Annamalai University Publications in Linguistics No. 5

INDIA AND HISTORICAL GRAMMAR

BY

M. B. EMENEAU

ANNAMALAI UNIVERSITY
ANNAMALAINAGAR
1965

INDIA AND HISTORICAL GRAMMAR

(Lecture on Diffusion and Evolution in Comparative Linguistics and Lecture on India and the Linguistic areas delivered as special lectures at the Linguistics Department of the Annamalai University in 1959)

BY

M. B. EMENEAU
Professor of Sanskrit and General Linguistics,
University of California, Berkeley

ANNAMALAI UNIVERSITY
ANNAMALAINAGAR
1965

Githa Press, Tiruchi-2.

PREFACE

Professor M. B. Emeneau came to South India in connection with the Summer School of Linguistics held at Coimbatore under the auspices of the Madras and Annamalai Universities, the Linguistics Society of India and the Deccan College, Poona. Annamalai University took advantage of the presence of this great savant for the inauguration of the Teaching programme of the Department of Dravidian Linguistics through delivering his three extension Lectures on the 19th and 20th of June 1959 under the presidency of the then Vice-Chancellor Sri T. M. Narayanasamy Pillai. These lectures emphasise the importance of diffusion in Comparative Linguistics, first, by considering the lexical borrowings and later, in the second lecture, the structural borrowings. In modern times areal Linguistics has come to compete with comparative Linguistics. In these lectures Professor Emeneau has elucidated and exemplified the methodology with the Dravidian materials.

In the modern set-up where the fundamental unity of India is emphasised, the conception of India as a Linguistic area cannot but be welcomed.

It is very unfortunate that these important contributions of Professor Emeneau could be published only in 1965. But the Linguistics Department is proud that it has this opportunity of publishing this valuable contribution of the Professor.

Annamalainagar,
16—12—1965.

T. P. MEENAKSHISUNDARAM.

CONTENTS

Diffusion and evolution in comparative linguistics 1

India and linguistic areas I 25

India and linguistic areas II 40

Bibliography 72

INTRODUCTION

These papers were delivered at Annamalai University on June 19th and 20th, 1959, as a series of three lectures, with some minor differences that need not appear in the published form. They formed part of the University Extension Lectures series, the first given under the auspices of the Department of Dravidian Linguistics.

The first paper had been delivered elsewhere, but it fitted so well in subject into the series that the opportunity was welcomed to deliver it again and to publish it as part of a larger effort to outline and define methods that need to be applied in historical and comparative linguistics when India is being considered as a whole-the whole being, as so often in scientific matters, the Indian subcontinent rather than the Republic of India.

For reasons that are clear to most scholars but that need to be restated often to the on-coming generations of linguistic students (as well as to laymen), the historical and comparative method that comes most easily to the mind and hand of the linguistic scholar is that of tracing genetic relationships of the evolutionary kind that can most conveniently be called 'continuants.' In the first paper there is discussion and exemplifying demonstration of the need to take into account explicitly also the other type of genetic relationship, 'borrowing' or 'diffusion.' The methodology set forth in the paper is minor in amount and perhaps not very profound; it does, however, have the virtue, I think, of explicitness The exemplification is based on studies in Dravidian, though a specialist in Indo-Aryan could have found many examples there.

Borrowing of lexical items has been so often treated that the obvious manifestations of this historical process have been neglected here, in spite of the rich variety of complication that the Indian scene presents. Only one such sub-type need be mentioned, which I have treated elsewhere (as have others) and which still needs treatment in much more detail than it has yet received, viz. borrowing between the three major Indian linguistic families. This and the simple intra-familial borrowings have been neglected here in favour of some Dravidian intra-familial borrowings which are difficult both to identify

and to work out in detail. Their difficulties whet the imagination and may perhaps in the long run lead to further statements of procedure.

The second paper attempts to carry further, both in methodology and in exemplification, the treatment of another type of borrowing, that of structural traits, which I initiated in an earlier paper, 'India as a linguistic area' (*Language* 32.3—16 [1956]). When such borrowings have gone on over centuries or millennia, they result in phenomena of the 'Sprachbund' or 'linguistic area' type. India seems to present a rich display of phenomena which point to this interpretation, and I have attempted in the present paper to throw the net wide geographically and to interpret the data in as detailed a way as possible, both to show what can be elicited from rather unsatisfactory data, and to point out the need of such studies by those who are better equipped for them than I am. The need for better descriptive data is always with us, and not least on the Indo-Iranian border, in spite of the valorous efforts of distinguished scholars in desperately difficult field conditions. Perhaps this study of mine merely points to one specific area where the linguistic scholars of India and Pakistan should be able to gather descriptive data with much greater ease and success than any Western scholar could hope or be expected to.

This introduction must close with an expression of my gratitude to the Annamalai University and its Department of Dravidian Linguistics for many things, and not least for the opportunity to present these papers before a lively audience and for their publication.

[Since the lectures were delivered, further study has added to and modified some of my views on details. It is impossible to make changes at all the points where I might like to do so. As a minimum, I have added a sentence and a footnote in square brackets.]

—o—

DIFFUSION AND EVOLUTION IN COMPARATIVE LINGUISTICS

In a paper on 'The Pattern of Progress in Romance Linguistics,' Professor Yakov Malkiel[1] viewed with approval "the recognition [in Romance linguistics] of [linguistic] diffusion as a major force, supplementary to rather than contrasting with evolution." This paper will attempt to sketch the history of the recognition of these two processes in the best-developed fields of comparative linguistics, and to adumbrate their application to Indian linguistics.

The early decades of Indo-European studies, in the first half of the 19th century, were marked by a struggle away from the methods best known in Voltaire's phraseology, in which "the consonants count for very little, and the vowels for nothing at all," a struggle toward something more systematic. In the early period scholars who were essentially learned, well-meaning amateurs, accepted, somewhat uncritically, any resemblance, either of sound or of meaning, and either partial or complete, as evidence of relationship between languages. The patent resemblance in sound and the identity of meaning between Latin *dies* and English *day*, and the somewhat less close resemblance in sound between Latin *duo* and English *two*, were both accepted

---

[1] *Romance Philology* 5·293 (1952). The present article is based on a lecture given in the Romance Philology Lectures series at the University of California, Berkeley, on November 3, 1955, with the title 'Dialectology and Comparative Linguistics.'

as valuable in showing relationship between these two languages. In fact, it probably invests the earliest investigations with more exactitude than they usually had, to attribute to them, in the light of modern doctrine, the view that one of these resemblances in sound is different in nature from the other. In essence, both pairs of words were regarded as equally good pieces of evidence in the early phase to which we are referring.

However, the need to make some systematic sense of much that was chaotic-seeming, especially in the Germanic material, produced a most notable advance in the work of scholars from Rask (1814) on through Jakob Grimm and many others. Whatever the actual statements of these men were, their method was to find as many examples as they could of sets of words which were related according to regular statements. Latin *duo* and English *two* were joined, as regards *d* and *t*, by many such pairs as *decem* and *ten*, *dens* (whose stem is *dent-*) and *tooth*, *pes* (whose stem is *ped-*) and *foot*, *id* and *it*, *quod* and *what*, *istud* and *that*, *edo* and *eat*, *sedeo* and *sit*, *ad* 'to' and *at*, *cor* (whose stem is *cord-*) and *heart*, *duco* 'I lead' and *tow*, *videre* 'to see' and *to wit*, Old Latin *dacruma* and *tear* (of the eye). In these pairs Latin *d* clearly is related to English *t*, and in all places in the word the same statement holds. Many more such statements of relationship were produced by these scholars. Not all of them were as simple as the one just exemplified, and in many of them it proved necessary to atomize the environment, to say that in different places in the word, e.g. initial, medial between vowels, in the neighbourhood of certain vowels or of certain consonants, in certain relationships to the accent, etc., etc., different sounds in the one language represent one and the same sound of the other. In the end it became necessary to recognize that the search for resemblances is only the initial trial, and often also the initial error, of the method. Latin *dies* and English *day* are not related

at all. Latin *dies* turns out to be related to English *tues-* in Tuesday, to Greek Zeús, and to Sanskrit *dya'us* 'sky, day.' English *day* may be related to another Sanskrit word for 'day', *a'har-*. Vice versa, Sanskrit *dīrgha'-*, meaning 'long', is demonstrably related sound by sound to English *long*, and Sanskrit *cakra'-*, meaning 'wheel', to English *wheel*; Greek *ὀμφή*, meaning 'voice', is related to English *song*. It is not possible to demonstrate these last pairs of words here; it is the subject-matter of an introductory course in Indo-European comparative grammar.

Other resemblances, to be sure, are found. Latin *dens* (whose stem is *dent-*) is related to English *dentist* and *dental*, Latin *decem* to English *decimal*, *decennial*, and *decimate*, Latin *cor* (whose stem is *cord-*) to English *cordial*, Latin *videre* to English *evident* and even to *video*. The resemblances here also can be dealt with in equally complicated sets of statements. In this particular instance Latin *d* is related (in spelling, at least) to English *d*.

What is the difference and the relationship between the two sets of statements? It is patent to all of us, as it was to the early scholars. The second kind of statement is interpreted historically as the result of borrowing. English scholars and litterateurs, when they have needed new words, have for many centuries usually gone to their knowledge of Latin and of Greek or to the dictionaries of these languages, and have borrowed wholesale. This is "linguistic diffusion," in one of its forms. The first kind of statement that we have referred to and exemplified is historically referable to some other process. In simplest terms, the two languages Latin and English are related by unbroken descent from a common ancestor, "which perhaps," as Sir William Jones said in 1786, "no longer exists." The common ancestor, which in our historical or prehistoric hypothesis is called Indo-European, is in fact no longer spoken anywhere and has not been for at

least 4500 or 5500 years. This is "linguistic evolution" or "genetic relationship."

When from our comparisons two parallel sets of statements emerge, as they do for the relationship between Latin and English, you will wonder how we decide which set represents "genetic relationship" and which "borrowing." The necessary assumption is as follows. It is first imperative to be sure that we have exhausted all the possibilities of different environments being responsible for the two correspondences. In our instance we must be sure that *t* and *d*, the two English correspondences for Latin *d*, do not occur in complementary environments, or, in other words, that *t* and *d* are in contrast in the same environment. Once we are sure that this is so, we then assume that of the two English correspondences, that which is closer to or even identical with the Latin sound has reached English by borrowing, and that which is less similar is genetically related. The English words with *d*, which is closer in sound to the Latin *d*, are borrowed from Latin, those with *t* are inherited from Indo-European. This, of course, is borne out by what we know of the literary and culture history involved, and the result in this instance can be reached directly from the culture history. Yet the methodological assumption is in the background of these studies, even if not explicitly stated, and when we lack older records and history, the method must be explicitly stated and used. Our early Indo-Europeanists, however, knew so much about the borrowings that had taken place from Latin and Greek into the modern European languages, and were so excited by the new discovery of genetic relationship that they concentrated on the latter, took the obvious borrowings for granted, and, while working on the assumption that I have stated, they did so more or less implicitly, so far as I can tell. I confess that I cannot say definitely that no Indo-Europeanist ever made the assumption explicit—the scholarly literature is too vast for me to examine even

superficially for this point. I must say, however, that it was not part of my training in comparative grammar under excellent teachers, and that I had to discover it for myself in circumstances that I shall mention later.

Indo - European studies went forward, then, with a progressive unfolding of genetic correspondences until in the 1870's numerous scholars came to the realization that their method looked for one and only one representation of an Indo-European sound in any one environment in one language. When an exception was found anywhere, a reason had to be found for it. Verner in 1872 had written "no exception without a rule" – certainly in opposition to the earlier practice of admitting frequent unmotivated exceptions. Leskien in 1876 made the classic statement: "phonetic laws have no exceptions." The older practitioners reacted violently, the younger ones stated their views tactlessly, and for almost two decades the debate raged acrimoniously, and, to be frank, has never ceased. The "young grammarians" or "Neogrammarians," including Leskien, Brugmann, Delbrück, Osthoff, and Paul, were forced to be very explicit about the "rules" which underlay the exceptions, including analogy and, of course, borrowing. If, as Leonard Bloomfield pointed out,[2] their slogan was unfortunately stated, since (as he said) "a linguistic change is not a 'law,' but a historical event, and ... the term 'exception,' accordingly, is irrelevant," and (as he also said) the more accurate statement is "phonemes change," yet for a quarter of a century or longer in the practice of the neogrammarians they forged a powerful instrument, or rather a tool-kit, and employed it skilfully. It does not do to concentrate attacks on their slogan and to ignore their procedure, to claim that they operated only with "phonetic laws" to the neglect of all else. These were, in fact, the scholars who recognized (if one may be facetious, but not overly so), that when the Romans

[2] *Language* 8·227 (1932).

perpetrated the rape of the Sabine women, these women's Sabellian kinsmen, including the red-headed Rufus, came to town to see their sisters, drove with them their cattle, and had refreshment in cookshops (perhaps kept by Sabines also)—and left behind them in the speech of Rome such words as *rūfus* 'read-headed,' *bōs* 'cow, bull,' and *popina* 'cookshop,' and others as well such as *scrōfa* 'sow (pig),' *vafer* 'sly,' and perhaps *cāseus* 'cheese' and *Caesar*. The rural dialectal origins of such Latin words were recognized by the neogrammarians. And it is only possible to recognize such borrowings when one's method has established genetic norms, i. e. phonetic laws from which the borrowings are deviations. But attacks have been made, which started legitimately enough with points at which the neogrammarians laid themselves open to attack, and then went on to establish an atmosphere, a climate of opinion, in which nothing that the neogrammarians did or taught could be right, in which in effect the baby was thrown out with the bath-water.

The methods employed by the neogrammarians in handling Indo-European were on the whole so successful that they were applied in other linguistic fields as well, notably in that of the Romance languages. I can speak only as an outsider, but I think I do not go far astray if I say that I gather that something went wrong. It turned out here, to put it over-simply, that there were too many exceptions to the phonetic laws. A series of scholars, culminating in, I suppose, Meyer-Lübke, persisted in the methods of the Indo-Europeanists. At the same time other scholars began to evolve other methods, particularly the collection and close study of dialect materials. The name of Gilliéron comes to mind and the doctrine associated with him, that "each word has its own history."

If I may proceed at once to the nub of the matter, the Indo-Europeanists and the Romance philologists turned out to be dealing with groups of languages to which the

same methods were applicable, but in differing degrees. The Indo-Europeanist studies a language family in which, by and large, at the time of the earliest records and for a considerable time previously, the languages most intensively studied were mutually unintelligible, there was no contact between them, and the parent language had become extinct long before without leaving a record. As between Latin and Greek, or Latin and Germanic, or Greek and Germanic, or any of these and Sanskrit, one set of correspondences only was found. It was only after the historical events which ended in Roman imperialism, that there appeared the conditions in which took place extensive borrowings which resulted in the double sets of correspondences with which our discussion opened. These were on the whole so transparent that the Indo-Europeanists could easily identify the borrowings and then ignore them. They were quite capable of dealing with less transparent situations also, as they did with that of Latin and its borrowings from neighboring rural dialects and languages. But in general, diffusion from language to language within the family had been so minor in extent that the old material usually studied offered few examples of it, and diffusion could be treated as the very minor factor that it was. The Romance languages were remarkably different. They were descended from the highly literate and prestigeful language of a great imperial power. Until the break which separated Balkan Romance from that of Western Europe, there had been no really effective barriers which could completely break off contacts between any two neighboring areas in the whole of this domain. New political entities of great cultural prestige arose, moreover, e. g. France, and their languages partook of this prestige. For almost 2000 years both evolution and diffusion took place, and the latter worked not only from numerous shifting centers but also from the literature of the old prestigeful parent language into its descendants. Here genetic correspondences of the kind seen in Indo-European studies were indeed to

be found and were asiduously sought by the older scholars. But diffusion had worked at so many levels over so long a period and the genetic correspondences came to be swamped by so many sets of diffusion correspondences, that it came to seem to many scholars that indeed "each word had its own history." And yet to the Indo-Europeanist it seems obvious that it is only by contrast with the genetic correspondences that the diffusion correspondences can be identified at all for what they are, that sound methods of distinguishing between the two must perforce have been worked out, and that the core of the methods is the assumption that has been made earlier in this paper. To quote Roland G. Kent:[3] "There are indeed so many factors which interfere with the operation of phonetic law, that some scholars now refuse to accept the law as a valid formulation; but their view is faulty, because they regard the interfering factors as fundamental and think phonetic law a mere interfering tendency, reversing the true relation." (Parenthetically, I would express my astonishment at the wilful blindness of such scholars as Bonfante. He once wrote:[4] "It would in fact be easy to show that for many of the so-called phonetic laws affecting ... Romance, ... the 'exceptions' are more numerous than the 'normal' forms. This is recognized [by one of my critics], who however makes the astonishing assertion that 'scientific principles do not in any way depend on the mere number of attestations.' One wonders on what they do depend." Bonfante knows very well on what canons they depend, and—to make it worse—later in the same article from which I quote,[5] he states that the areal principles of the neolinguists do not proceed by counting the noses of the majority and takes to task the neogrammarians for paying attention to relative numbers. He cannot have it both ways.)

---

[3] *The Sounds of Latin*, §5, n. 2.

[4] *Language* 23·345, n. 5 (1947).

[5] *Ibid.* 366.

No scholar in the Romance field has, I suppose, really abandoned the method that was central for the Indo-European scholar. In varying degrees, however, different Romance philologists have shown dissatisfaction with the methods of the neogrammarians as applied to the Romance languages. The French dialectologists, to be sure, have always exhibited their dialect material against a backcloth of regular correspondences, and have recognized that the genetic correspondences are disturbed and changed by diffusion and other factors as well. Gilliéron spoke justly of "l'implacable régime de la phonétique,"[6] i. e. inescapable phonetic laws. Others, the neolinguistic scholars, have in theory disowned phonetic law completely and diffusion as well, as principles that can be applied, and have vilified the neogrammarians with most intemperate language. They have even attempted to dismiss phonetic law from the Indo-European field and to substitute there the areal principles which they have developed for the Romance field. Within what is essentially a great unbroken dialect area in which diffusion can operate very freely, the areal principles can probably be applied with much profit, as they most certainly can *not* over most of the Indo-European field, in which mutually unintelligible languages are found, without effective contacts and lines of communication at the time of our earliest records. This had perhaps not been effectively pointed out by the critics of the neolinguistic group until L. R. Palmer did so in 1954 in his book *The Latin Language.*[7] Few other scholars have been so intransigent as some representatives of the neolinguistic group, but many others have been willing to belabor the neogrammarians with one stick or another. They have been willing to argue that "because the results of borrowing ... are distinguishable by their phonetic

---

[6] *Généalogie des mots qui désignent l'abeille*, 14; I know this from Bloomfield, *Language* 4·288, n. 6 (1928).

[7] pp. 26–30.

deviations, therefore the assumption of regular phonetic change is invalid"[8]—a patent fallacy. And yet it has frequently been pointed out that when they have made contributions other than theoretical to the discipline, it has been by methods that are not to be distinguished from those of the neogrammarians whom they traduce. In other words, whatever they may claim, phonetic law is there as a hidden assumption.

Professor Malkiel's phrase with which I started represents the only attitude that can be fruitful. Some Indo-Europeanists, especially in Europe, are still in the throes of theoretical reaction against the neogrammarians and may still take some time to regain balance and to cease from useless polemics. In practice they follow the well-established methods of the discipline. In the United States of America the fashionable European revolt against the neogrammarians hardly touched native scholars. The reason is interesting to note. Most of the training of young scholars in this field was supplied by a very long-lived pupil of the neogrammarians themselves, Brugmann and Leskien, namely, Carl Darling Buck, who died early in 1955.[9] From 1892 on, Buck and his pupils trained nearly all the Americans who have worked in Indo-European. The European contagion did not reach these younger men, especially since the habit of going to Germany for training ceased decades ago. By the same token, it is unfortunate that younger American Indo-Europeanists were not directly exposed to the European interest in diffusionary studies. They have been long in appreciating the profit that can result from the amalgam of evolution and diffusion that Professor Malkiel wrote of (and that, incidentally, is there in Bloomfield's *Language*, chapter 19, for all to read—there is too much in Bloomfield to absorb easily at first reading). At the same time, however,

[8]Bloomfield, *Language* 8·229 (1932).

[9]Cf. George S. Lane, *Language* 31·181 ff. (1955).

one must point out that the awakening has certainly begun. E.g. in the investigation of one of the matters where advances have been made by the new fusion, that of the relationship between the Italic languages of the Indo-European family, contributions have been made by an American scholar, Madison Beeler.

I have been forced to think of these matters by my attempts to do comparative grammar in a field far removed from Indo-European or Romance philology. The Dravidian family of languages formed the subject of my linguistic fieldwork in India in the three years 1935-38, My work was primarily descriptive and centered on four languages spoken by non-literate, essentially backwoods groups, three in South India and one in Central India. Even while I was engaged in fieldwork, questions of linguistic diffusion forced themselves on my notice, since no one of these languages is spoken in a vacuum and borrowing has continually taken place between them and their neighbors who speak different languages.

The question of genetic relationship also could hardly fail to interest a person who had been trained in Indo-European scholarship by such scholars as Franklin Edgerton and Edgar H. Sturtevant. That the Dravidian languages make up a family separate from all other families has been clear, even obviously clear, since as early as 1816, nearly a century and a half ago. Many phonetic correspondences have been obvious on superficial inspection, and scholars have often made the observation that the genetic relationship between the languages is about as close as that between the Romance languages, rather than like that between the old Indo-European languages. In other words, the degree of phonetic divergence is not very great. Questions of chronology and rates of change are at once suggested. but the matter cannot be pursued here, beyond the suggestion that there has been a slower rate of change in general in the Dravidian family than in the Indo-European.

In my handling of my fieldnotes since 1938, I included almost from the beginning the first step of comparative work, namely, the search for etymologies. By what I have just said about the degree of phonological differentiation, it can be guessed that etymologies were not too difficult to find. But, as soon as uncertainties arose, and they arose almost at once, it became necessary to state explicitly, even though tentatively, phonetic correspondences of the neogrammarian's sort, so that etyma could be correctly identified. Complexities soon appeared involving different correspondences in different positions within the word. Occasionally analogic disturbances could be identified. In fact, the procedure went on typical Indo-European lines.

In my concern over getting some of my field material completely analyzed and published before my retirement and death, I decided finally to make an all-out effort first on that language for which I had least material, namely, the Kolami language spoken in Central India by a community of about 40,000 people, and the book was published in 1955.[10] I had spent only six weeks on this language in the field and had concentrated there on a correct phonetic recording and a definitive phonemic analysis of everything that I took down, even if thereby the bulk of material should be reduced. The procedure was abundantly justified in all my later study, descriptive or comparative, since it enabled me to make statements of both kinds with a high degree of certainty. In completing my comparative work on this language I made the collection of etymologies as exhaustive as I could, recording from all the available accounts of the other 18 or 19 Dravidian languages all the etyma that I could identify. I also of course worked out the phonetic correspondences as completely as could be done. To my

[10] *Kolami, a Dravidian Language* (University of California Publications in Linguistics, vol. 12)

discomfiture it turned out that for numerous of the proto-Dravidian phonemes, Kolami had two correspondences, even after all possible positional environments had been taken into account. This reminds one, of course, of the situation as between English and Latin which was discussed earlier, or to project the same data onto the larger scale, the resultant two relations between English and proto-Indo-European, in one of which English shows a continuant from Indo-European, but in the other Latin shows a continuant from Indo-European, and English then a borrowing from Latin. In the Dravidian instance the assumption that I presented earlier was applied in the generalized form as follows: when for the same phoneme of the proto-language, in the same phonemic context, two different phonemes are found in language A corresponding to one phoneme in language B (or to two phonemes whose distribution is explainable by contextual conditioning), then that phoneme of A which is the more dissimilar to the phoneme of B is the straight-line development, i.e. the continuant,[11] from the proto-language and the other is found in borrowings from B. In general, this applies to languages which are in such a situation that contact of some sort allows or allowed borrowing at a postulated period. It is notable that relative numbers of examples do not enter into consideration at all (cf. my earlier strictures on Bonfante who apparently tried to have it both ways in his use of statistics).

In addition to this procedural principle others are to be invoked which need not be set forth here, though one may be mentioned, Details of morphology and of meaning, as well as of phonology, must be brought into play in identifying borrowings, and roughly the same principle of greater and less similarity applies in handling all three.

---

[11]Bloomfield, *Language* 393, 404, used the word 'continuant' as a quasi-technical term, as my colleague, Professor C. Douglas Chrétien, has reminded me.

It turns out that Kolami has borrowed extensively from Telugu. Kolami is spoken in an area abutting on the northwest corner of Telugu territory. There has been expansion of Telugu speakers of various occupational castes, especially farmers and shopkeepers, into the territory where the Kolams are found, and there has undoubtedly been a long history of this trend as well as of shifting of the areas inhabited by Kolams. At any rate, there is here a complex history, partly still unguessed, of interrelation between speakers of the two languages, and there has been much bilingualism of the sort in which the members of a minority community learn more or less of the language of the majority community which possesses prestige.

One example of the type of data will suffice. PDr. *ṟ : Old Ta. Ma. To. Old Ka. Old Te. ṟ, Ko. Koḍ. *r*, Go. r/*rr*, Kui *j*, Kol.-Nk. *d*, Pa. *d*, (NE) *ḍ*, Gaḍba *y* ; in modern Ta. Ka. Te. ṟ and r fall together, as they do in Ko. Koḍ. Ta. *maṟa*, Te. *maṟacu*, Kol. *mad-* 'forget'; Ta. *kuṟaṅku*, Te. *kuṟuvu*, Kol. *kudug*, (stem *kudg-*) 'thigh'; to be contrasted with Ta. *āṟu* 'heal,' Ka. Te. *āṟu* 'become dry, heal,' Kol, *a·r-* 'become dry'; Ta, *maṟu* 'another, next,' To. *maṟ* 'next in time,' Ka. *maṟu*, *maṟa* 'other, next, again,' Te. *maṟu*. 'other, next,' *maṟi* 'and, again, then,' Kol. *mari* 'thereafter, then (in past or future)'. In the last two instances borrowing is to be inferred; in the last one, morphology as well as phonology provides evidence.

One corollary of some interest is the following: once a high incidence of borrowing by A from B has been established, it becomes impossible in a set of etyma drawn from A, B, and one or more other independent languages (C, etc.), to determine whether A's item is a continuant from the proto-language or a borrowing from B, if on establishment of the relevant phonological

correspondences no distinguishing features mark off A from B, C, etc. in this set of etyma. One very clear example is Ta. Ka. Te. *allu*, Kol. *al-*, all meaning 'to weave.' Even though the Kolams are not weavers and the word might be a borrowing from Telugu, it might just as well on phonological grounds be inherited from proto-Dravidian. And there are numerous other examples in which both the linguistic data and the non-linguistic factors are entirely indeterminate.

In my examination of the vocabulary collected by myself and two other smaller vocabularies that have since become available, 720 items were found to have Dravidian etymologies. Of these 166 are certainly borrowings from Telugu, i. e. about 23 percent of this part of the vocabulary. 395 items, i. e. 55 percent, are certainly not borrowings from Telugu. The remaining 159 items, i. e. 22 percent, belong to the indeterminate catagory. The matter might be let rest at this point. A possible way, however, of treating these last items in percentages is to assume that in them the proportion of borrowings and non-borrowings from Telugu is the same as in the other part of the vocabulary. When this bit of arithmetic is worked out, the resulting approximate percentages of borrowings and non-borrowings from Telugu into Kolami in the whole vocabulary of Dravidian origin are 30 and 70 percent.

We are then in this part of the Dravidian field, namely, Central India with its numerous minor languages under the shadow of several major ones, dealing with a situation which is much more like that of the Romance languages than that of the old Indo-European ones. Techniques of dealing with diffusion must be adopted or evolved to suit the requirements of the field. Not least important will be the determination of historical and social factors. The political history is fairly clear, but it may in the long run be less important than the

socio-economic structure with its graded castes or caste-like communities, between which complicated prestige factors and combinations of prestige factors operate to determine the directions and the degrees of diffusion. Nearly everything still remains to be done in this geographical area, beginning with adequate descriptions. I should say in passing that the picture is a great deal more complicated than I have just indicated, since there are languages of three families, Dravidian, Indo-Aryan, and Munda, found in the area in more or less close symbiosis. The Kolami speakers with whom I worked in Wardha district were in close contact with speakers of Indo-Aryan Marathi and Dravidian Gondi; the vocabulary that I gathered from them contained 35 percent of words with Marathi etymologies. In one area Dravidian Kuru*kh*, Munda Korowa, and a local variety of Indo-Aryan Hindi are all spoken. Neither of these situations is of an unusual type. One of the resultant phenomena is special local vocabularies which belong to all the languages of the locality involved but which in the present state of our knowledge cannot be etymologized as belonging to any one of the families apart from the locality. There are undoubtedly also other locally circumscribed phenomena-phonological, morphological, and syntactic. This type of linguistic symbiosis as a part of the social symbiosis of the caste structure and of its marginal manifestations in various areas of India, has hardly been more than envisaged so far. Studies of the linguistic phenomena of diffusion in this type of situation promise, however, to yield fascinating results in the future.

Lack of data prevents more than guesses about the linguistic situation further south. The four major languages, Tamil, Malayalam, Kannaḍa, Telugu, all have ancient literatures and tens of millions of speakers (the Dravidian family as a whole has approximately 90 million speakers). These languages clearly have

local dialects, but we know remarkably little about them in modern or any other terms. We know little too about the nature of the boundaries between the four languages. Are they like that between German and French, or like that between French and Italian? It seems fairly clear that Telugu is well marked off from both its neighbors, Tamil and Kannaḍa. Malayalam and Tamil are very closely related, in fact, have been separate languages for less than a millennium. They are separated along most of their north-to-south boundary by a heavily jungled, and in consequence very sparsely settled, mountain range. across which there is in effect no communication. At the south end of this line the mountains level out and the linguistic boundary, we are well informed,[12] is a fan of isoglosses like many of the boundaries within the Romance or the Germanic area. There may be a similar boundary at the Palghat gap further north, but we are not informed. Nor have we any information about the nature of the boundary between Tamil and Kannaḍa, as to whether there is a clean break or whether border dialects are to be found. Only dialect atlas work will answer questions of this kind. Again there is the problem of well-marked symbiotic dialects whose boundaries are the social ones of caste organization. We know hardly anything of this, and any dialect atlas will have to provide the needed data. After the descriptive work is done, it will be possible to study the various directions of diffusion. Something is already known of the diffusions between the major languages that have been contained at various periods within the larger shifting political entities of the south. E.g., the Vijayanagar kingdom from 1336 to 1565 A. D. contained large blocks of Telugu, Kannaḍa, and Tamil speakers, and it was only one of a succession of such kingdoms. Certainly there

---

[12]So I would interpret the information about these dialects given by C. R. Sankaran and A. C. Sekhar, *BDCRI* 7·220–4 (1946).

has been much borrowing between Telugu and Kannaḍa, and between each of these and Tamil,—all these borrowings being identifiable by the usual methodological principles; but we certainly need further detailed work on the philological facts—when are borrowings first attested, and what are the correlations with the political history?

Another interesting area is the plateau of the Nilgiri hills, in which I did extensive work on two languages, Toda and Kota, and a very little on a third, Badaga. Here in a territory approximately 40 by 20 miles, these three communities, which should really be considered castes in their overall relationships, live in symbiosis with one another, and the three castes speak mutually unintelligible languages. The high plateau was virtually isolated from the plains below by lofty jungled precipices until about 100 years ago. The Todas were certainly there in 1117 A. D., when a general from the Kannaḍa country made an excursion into the hills and recorded the presence there of Todas, and I have recently [18] attempted to prove on linguistic grounds that the Toda language is connected with pre-Tamil and have suggested that there is at least a possibility that the Todas have been in the Nilgiris for the whole period of their separation from Tamil speakers. Of the history of the Kotas we know nothing, but they and the Todas agree that both were created in the hills where they are now — for what such a tradition is worth. We have only the most legendary of history for the Badagas, but they were certainly there in 1603, when the Portuguese priest Finicio visited the hills and found the Todas and Badagas in about the same relationship that they now share. It is fairly clear on linguistic evidence that the Badagas, who speak a dialect of Kannaḍa, arrived in the Nilgiris from the Kannaḍa country after the Kannaḍa language went through the change of Old Kannaḍa

---

[18] 'Toda, a Dravidian Language,' *Transactions of the Philological Society* 1957. 15-66.

initial *p–* to Middle and Modern Kannaḍa initial *h–*, since Badaga too has initial *h–* in the relevant words. The change began in Kannaḍa inscriptions in the 10th century A. D. Neither Toda nor Kota shows the change. The appearance of the Badagas in the Nilgiris must probably be placed even later than the 12th century, since the Badaga dialect, like Kannaḍa, has only one *r*-phoneme. In Kannaḍa inscriptions the alveolar *r* and the post-dental *r* begin to fall together as early as the late 12th century and are evidently not distinguished at all in the vernaculars from the 13th century on. We need a good description of the Badaga language before further linguistic attempts can be made on its history.

One of the more important problems of Nilgiri linguistics, with all its bearings on the prehistory of the peoples of the Nilgiris, is the relationship between Toda and Kota. Certainly they show many features in common, notably the loss of most unaccented vowels, and a certain amount of shared vocabulary that either is not found elsewhere (or at least has not been identified elsewhere), or differs in morphology or meaning from related vocabulary items elsewhere. A few examples will suffice. The Nilgiri word for 'lung', Ko. *podn*, To. *pï̈θ* (entry 3687), has been found nowhere else. The morphemes meaning 'to embroider' are Ko. *oṛk–*. To. *wï̈ṛk–* (entry 851), and have been found nowhere else with this meaning; related words elsewhere are Ta. *oṟukku* 'to draw out, as gold thread', Te. *oḍuku*, *vaḍuku* 'to spin (transitive)" Tu. *nūl-oḍu* 'a spindle', Kui *o̱sa* 'to sew, weave', Kur, *ōjnā* 'to spin, twist', Malto *ò̱je* 'to twist'. Beside Ko. *oy̤*ḷ, To. *wa̱sy*. both meaning 'waterfall' (entry 840), the other languages have only words with the more generalized meaning 'to flow' (Ma *oliyuka*. *olikka*, *ōluka*, Te. *oluku*), or in Tamil *oliyal* 'river', which is the nearest item in meaning to the Nilgiri words, but is morphologically somewhat different. In spite of many such coincidences between the two

languages, there are numerous differences also. For a considered opinion as between genetic and diffusional relationships between the two languages, the necessary detailed work must be done first.

Borrowings into both languages from the neighbouring Badaga are clear already, even though we know all too little in detail about Badaga. Toda has borrowed numerous words. These, curiously, tend to appear in the songs, alongside of the continuant forms from proto-Dravidian found in Toda conversational language. E.g., in ordinary Toda the word for 'snake' is *po·b*, cognate closely with Ta. *pāmpu*, Te. *pāmu*, less closely with Old Kannaḍa *pāvu*, later Kannaḍa and Badaga *hāvu*. The latter appears in a Toda song as *o·f*.

A more striking instance is the common Kota verb stem meaning 'to go.' All the South Dravidian languages, including Toda, have a stem *pō-* (or its descendants) with various irregularly related allomorphs such as *pōk-*, *pōy-*, *pōv-*, and a past stem *pō<u>n</u>-* (entry 3734). Modern Kannaḍa of course has forms with initial *h-*. The Kota forms show the usual irregularities of this verb, but have no initial consonant : *o·g-*, *o·y-*, past stem *o·n-*. Absence of initial *p-* must be related somehow with the Badaga verb with initial *h-*. Only a few of the forms of the verb show close resemblance otherwise to the Badaga forms ; for the rest, the Kota verb structure is not close to that of Badaga. It must have been a very few similar forms (e.g. imperative, Badaga *ho·gu*, Kota **po·g > o·g*) that provided the starting point for the loss of *p-* in the reminder of the Kota verb. It is to be noted that the Kota verb *po·k-* 'to spend time,' which is historically a causative of the verb meaning 'to go' has initial *p-*. Kannaḍa, and presumably Badaga, do not have this form. On the other hand, Tamil alone of all the Dravidian languages has a causative of this form with this meaning (*pōkku*). It is certainly possible, though not necessary

to assume that the Kotas have borrowed the word from Tamil speakers; another possible assumption would be that Kota, like Toda, is most closely connected with pre-Tamil, and in that case Kota might have inherited *po·k-*.

The conspicuous difference between the older Nilgiri languages and Badaga in the matter of initial *p-* seems to have led to other contaminations of the kind seen in the Kota word 'to go.' Both Kota and Toda have several notable examples, in which the situation is complex. The South Dravidian languages have a verb whose meanings fall within the range 'to carry, sustain, endure, undertake, obey' (enty 3729). The reconstructed stem coincides with that found in Tamil, **poṟu*. *ṟ* is the alveolar *ṟ* that was mentioned above, which appears in Old Tamil, Old Kannaḍa, Malayalam, and Toda as *ṟ*, a phoneme distinct in all these languages from *r*; in Kota, as in Modern Kannaḍa and Badaga, the two phonemes have fallen together. The Nilgiri languages and Kannaḍa have a past stem which derives from **poṟṟ-*. In Old Kannaḍa already it is *pott-* with dental *t*. In both Kota and Toda **ṟṟ* becomes *ṯ*, which is alveolar and phonemically distinct from dental *t*. This verb, then, in Old Kannaḍa has stems *poṟ-*, *pott-*, in Modern Kannaḍa and Badaga *hor-*, *hott-*. Kota has the continuant forms that are to be expected in accordance with the genetic developments: *por-*, *poṯ-* 'to carry, obey.' So has Toda: *pīṟ-*, *pīṯ-* 'to carry.' In addition Kota has *or-*, *oṯ-* 'to carry, undertake, obey,' and Toda has *wīṟ-*, *wīt-* 'to undertake, hold, carry.' These forms without *p-* are due to Badaga influence, but it is to be noted that Kota keeps the genetically expected alveolar *ṯ*, and Toda in contrast does not have the genetically expected alveolar *ṯ*, but does have the genetically expected alveolar *ṟ*. Kota has been influenced by Badaga only in the loss of *p-*; Toda has both this feature of Badaga and the dental *t* of Badaga in the past stem, but has retained from its own first stem

pīr- the *r̲* that disappeared as a separate phoneme in Kannaḍa already in the 12th and 13th centuries A.D. We have here the two languages borrowing independently from Badaga, but the forms show the influence of diffusion only partially. Or, in other words, as was said above of Kolami and Telugu, the minority community learns 'more or less' of the language of the majority community.

In Toda there are a few other striking examples. One will suffice, the name of the god of the dead, *ö·n*. The set of etyma (entry 3635)[14] is: Ta. *pēy* 'devil,' *pēyan̲* 'demoniac, madman,' *pēycci*, *pēcci* 'demoness, woman under possession of a demon,' Ma. *pē* 'demon' (fem. *pēcci*), *pēna* 'ghost, spirit,' Ko. *pe·y* 'demon,' *pe·n*, *pe·nm* 'possession of woman by spirit of dead,' To. *ö·n* 'god of the dead,' Ka. *pētu*, *hēde*, *pēṅkuṇi*, *pēṅkuḷi* 'demon,' Tu. *pēyi* 'demon,' Go. *pēn* 'a god,' Kui *pēnn* 'a god, a spirit.' These appear all the way from Malayalam on the Malabar coast to Kui at the northeast corner of the Dravidian territory. A proto-Dravidian morpheme is indicated, probably the old word for 'god.' Toda and Kota share a meaning connection with the dead; it is still uncertain whether this is shared innovation, or shared retention of a feature of proto-Dravidian or proto-South Dravidian. We must ignore here all other problems than the Toda lack of initial *p-*. It looks like Kannada or Badaga influence. For neither language, however, is a form **hēnu* or **pēnu* on record, and it is a problem why without a model in Badaga the Todas should have lost initial *p-* in this word. There should be added another instance that surely has a bearing on this word. 'Corpse' (entry 3420) in general South Dravidian is **piṇam*, which appears in Kannaḍa as *peṇa*, and then *heṇa*. In Toda it is not the word for 'corpse' but

[14] Already presented for another purpose in *American Anthropologist* 55.453–4 (1953).

for 'the dead'; *ī̤no·ṛ* is 'the world of the deəd' (*no·ṛ* 'place'), *īṉṭöw* is another name for *ö·n*, the god of the dead (*töw* 'god'). The Toda words, considering the mythology and eschatology of this caste, are likely to be old in Toda, and not merely borrowings from Badaga. The loss of *p-* in *ö·n* and in *īṉ* is undoubtedly connected. But why did the Kannaḍa-Badaga loss of *p-* strike these words particularly, even when *p-* was normally retained? Are there perhaps factors of linguistic taboo operating in connection with words for the dead, that allowed partial distortion of the old Toda words through the diffusion of the conspicuous difference in respect to *p-* between Badaga and the Nilgiri languages? Close examination of the Toda prayers and songs will probably yield a little more evidence of the same kind, but there has not yet been complete utilization of this material. Diffusion has surely operated in the Nilgiris, but sometimes in peculiar ways.

Even though so little is yet known about Badaga in detail, I have already[15] pointed out what in all probability is an instance of Toda-Kota influence on Badaga. The verb **ta-*, **tar-*, **tār-*, past stem **tant-*, in a number of the southern languages, viz. Ta.-Ma., Ko., To., Koḍ., means 'to give to a 1st or 2nd person,' in contrast with **koṭu-*, **koṭṭ-* 'to give to a 3rd person.' In Kannada the continuants of **ta-*, etc. mean 'to bring,' but the Badaga forms have the same meaning as those of Toda and Kota. Unless we are to posit, improbably, that Badaga retains a pre-Kannada meaning, this is an instance of borrowing by Badaga from the old Nilgiri languages.

To conclude, in any study of Dravidian comparative grammar and etymology both evolution and diffusion must be taken into account. The territory strikes one as a large dialect area of the Romance nature, with numerous

---

[15] 'The Dravidian Verbs "come" and "give", *Language* 21.184–213 (1945) esp. 190.

boundaries marked by thick clustering of isoglosses, but also with many, many isoglosses that enclose more than one language and with many instances of multiple phonetic correspondences. These latter features of the family are only recognizable for what they are, after phonetic correspondences are looked for according to the methods of the neogrammarians. Then it becomes necessary to deal seriously with diffusional problems.

In this paper I have dealt only with the application of evolution and diffusion to Dravidian material. It will have been obvious that problems of the same kind exist in the historical interpretation of the data of Middle and Modern Indo-Aryan. The application to them of these two types of treatment yields statements about continuants of Sanskrit forms *(tadbhavas)*, borrowings from Sanskrit into Modern Indo-Aryan, borrowings between various Middle Indo-Aryan languages, borrowings between various Modern Indo-Aryan languages, and the like. Again, the Indo-Aryan languages in their spatial and temporal relations resemble the Romance languages. I shall not even attempt to exemplify this. Specialists in the Indo-Aryan field have done so and of course will do so in the future. I can only recommend to them that an explicit recognition of these different types of historical relationships may be of some profit in solutions of tricky problems, and with this cease.

INDIA AND LINGUISTIC AREAS[1]

I

It has long been a question of dispute among linguistic scholars whether and how much languages influence one another's structures. That they influence one another's lexica is beyond doubt, and will not be discussed in this paper. The question at issue has to do, not primarily with lexical items, but with phonology, morphology, and syntax. In so far as the question of the lexicon is in many languages inextricably tied up with derivational morphology, the unquestioning acceptance by scholars of lexical borrowings between languages has involved for some languages an acceptance also of the borrowing of derivational morphemes. Edward Sapir firmly as he held the view that languages hardly, if at all influence one another in their "intimate" structures, yet was unable to avoid discussing the fact that English has borrowed from the classical languages and (or through) French such derivational morphemes as *-ess*, *-ize*, *-able*, and that they are of great productivity in the language, occurring with other nuclear forms than those derived from Latin and Greek (or French) with which they were originally borrowed. These, he insisted, were merely additions to the old stock of affixes and hardly differed from borrowings of words; he found them to be superficial changes, and not "intimate" in the sense he accepted.

---

[1]I must record my gratitude to Miss Catherine Callaghan, who did a great deal of preparatory bibliographical work for this paper, and to the Research Committee of the University of California, Berkeley, for financial aid that made her help possible. I should also record the enlightenment that resulted from numerous conversations on the subject with my colleagues.

In an earlier paper, 'India as a linguistic area' (*Language* 32·3–16 [1956]), I discussed this attitude of Sapir's as part of a controversy which was carried on for a number of years especially between him and Boas as protagonists, usually in the non-explicit, anonymous exchanges that are seemly between pupil and teacher. Boas over many years published longer or shorter statements of his findings on the inter-influences between North American Indian languages, especially of the northwest coast, certainly in phonology and in all probability in morphology and syntax.[2] The absence of time-depth in our knowledge of these languages made it particularly hard to offer convincing demonstrations, and the demonstration for these languages still remains essentially where Boas was forced to leave it, and will remain there until very intensive descriptive and comparative work has provided the remedy for lack of easily handled historical data.

In my paper I demonstrated that India with its three major language families offers convincing evidence of inter-influence between them, and that besides phonetic and phonological traits which seem clearly to have been borrowed from one family into another, there are also morphological and syntactic features that are explicable in the same way. I used of India the phrase 'linguistic area,' which I intended to be taken as a technical term

---

[2]I gave as partial bibliography his introduction to *Handbook of American Indian Languages* 1·47–53 (Bureau of American Ethnology, Bulletin 40, part I: Washington. D. C., 1911), *American Anthropologist* 22·367–76 (1920), and *Language* 5·1–7 (1929). I should also have included the important longish statement in *International Journal of American Linguistics* 1·1–8 (1917). Sapir's opposing point of view was set forth in 1921 in his book *Language*, pp. 205–20. There he did not name Boas, nor did Boas name him in his 1929 article.

equivalent to the Germen *Sprachbund*. The definition that I gave (p. 16, n. 28) is as follows: 'an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families.' Weinreich, in *Word* 14·379 (1958), discusses this problem of the *Sprachbund* and finds fault with all terms that have been suggested for the phenomenon. In fn. 6 he suggests 'convergence area' as perhaps "more specifically meaningful" than my term. I would argue that convergence in the sense applicable here is the effect of the process of diffusion and that 'diffusion area' might be an equivalent term of specific meaningfulness. I would also however argue that a dialect area, according to the historical interpretation that follows the doctrine of Johannes Schmidt, is also a diffusion area or convergence area. These terms are still not specific enough, since they refer to a historical process and the result of that process, and since the same process produces the same result whether between dialects of a language, languages of a family, or languages of several families (so at least I am concerned to demonstrate). A more specific term for 'linguistic area' in the sense in which I have defined it would be 'multifamilial convergence (or diffusion) area,' as opposed to 'multilingual convergence (or diffusion) area' and 'multidialectal convergence (or diffusion) area.' These terms (or equivalents based on Greek) are however cumbersome, and I still prefer the simpler (if still unsatisfactory) arbitrary term and the exact definition I have given it.

When a map is used to illustrate the matter, any trait that is being discussed may be placed on the map with the aid of a line encircling the languages that have the trait, and this line, like that used in dialect studies, can be called an isogloss. The use of this term and illustrative technique involves, of course, the assumption that we are using in this matter the same assumptions that are made

in dialect studies, concerning spread of a trait in the manner of a Schmidtian 'Welle,' the prestige of one community over another as the chief social factor causing adoption of traits, and bilingualism as a necessary linguistic factor in the process especially where different languages are involved, but also even when it is a matter only of dialects.

On the other hand, the boundaries between related dialects and between closely related languages emerge as thick bundles of isoglosses. Prima facie, one might expect that a linguistic area in the present sense might be delimited in the same way. Unfortunately, I know of no demonstration of such a bundling of isoglosses. In fact, when in the earlier paper I treated India as a linguistic area, I made no attempt to demonstrate a bundling of isoglosses, but I rather discussed a number of traits that cross family boundaries in India and I was concerned as regards one or two of them to demonstrate the 'Indianization' of Indo-Aryan, i. e. to demonstrate that Indo-Aryan at various periods shows traits that originated in Dravidian and spread over more or less wide Indo-Aryan territories. The affricate pronunciation of earlier palatals is a case of less wide distribution, since it concerns chiefly Marathi, southern Oriya, Telugu, and northern Kannaḍa, and probably some minor languages of central India; i. e. this isogloss delimits a central Indian linguistic area involving certainly two, and probably three, families. Much wider distribution, which can loosely be called pan-Indic, is seen in the isoglosses that represent the occurence of retroflexes and the occurrence of non-finite verb forms.

The former isogloss as discussed in the earlier paper represents the presence on the phonological level of retroflex consonants in practically all the languages, whether they belong to Indo-Aryan, Dravidian, or Munda, and no matter to what chronological period the evidence belongs (p. 7). The discussion of the necessary historical statements need not be repeated here. What is of present

interest, however, is the further extension of this isogloss (that was not made in the earlier paper) to include some of the Iranian languages.

The material on some of these languages is scanty. For Balochi it is somewhat uncertain what the true state of affairs is. The LSI sketch (X, 336) states that the western dialects, which represent " the oldest and most original form of the language, " have *ṭ*, *ḍ*, *ṛ* in Indo-Aryan loanwords ; it certainly implies (p. 337) that the eastern dialects also have them. On the other hand, Gilbertson (p. 7), without distinguishing dialects for this feature, says that the Indo-Aryan retroflexes are replaced by Balochi dentals —" but this is not a hard and fast rule by any means. " He finds however that loanwords from English have retroflexes rather than dentals to represent the English alveolars. Whatever the phonemic status of retroflexes may be in Balochi, it is clear that only borrowed words have them in any dialect or idiolect.

The Pashto of Peshawar, according to Morgenstierne (1926, p. 12), "has been largely Indianized in its phonetic system." Retroflexes occur in this dialect, and also in that of Kandahar as described by Penzl, in loanwords from Indo-Aryan, " also in many words of uncertain origin " (so Morgenstierne) which also contain phonemes that are not Indo-Aryan, and certainly in some Iranian words as well, e. g. that for ' porcupine ' which Morgenstierne (p. 22) gives as *ṡkūṛ* and Penzl as *shkunn* (*nn* being his writing for the retroflex nasal).

In Ormuri and Parachi, as described by Morgenstierne (1929. pp. 16, 41, 331, 333, 335), retroflexes occur, and not only in words borrowed from Indo-Aryan. For the languages further north, those of the Pamir and others, all called by the LSI Ghalchah, the evidence is good on some of the languages, but very scanty on others. Yidhga, according to Morgenstierne (1938), has *ṭ*, *ḍ*, *ṛ*, and *ṇ* ; *ṭ*

and ḍ seem to be found only in loanwords. Morgenstierne's account of Sanglechi–Ishkashmi (1938) has ṭ, ḍ, ṇ, ḷ, with dialectal differences of occurrence. The Wakhi material given by Morgenstierne (1938) and Lorimer (1958) has some retroflexes, but there is doubt about their status. Persian and the western Iranian languages do not have retroflexes.

The isogloss then runs roughly north and south through Afghanistan and Baluchistan, but in the present state of our descriptive knowledge it is impossible to draw it with any accuracy. It is impossible, moreover, to suggest in any but the most general way how the retroflexes have developed in the Iranian words in which they are found in such languages as Pashto and Yidgha. Bilingualism, involving Indo-Aryan languages, must be the answer, but no historical and social details can be given to assist in elucidating the process. To the question why the undoubted bilingualism between Balochi and Brahui speakers has not resulted in Balochi's acquiring retroflexes, the answer is undoubtedly Bray's, that this bilingualism is on the whole a one-way affair, in which the Baloches seldom learn Brahui. Brahui loanwords have been identified in Balochi, but the conditions have not been present for 'intimate' Brahui speech-habits to be adopted in Balochi. The time during which there has been contact between Brahui and Balochi, moreover, is comparatively short, as will be shown later. And yet one wonders if the last word has been said in this matter. [The matter is further treated in my *Brahui and Dravidian Comparative Grammar*, Chapter IV.]

The second Pan-Indic isogloss mentioned above is that concerned with the use of non-finite verb forms (gerunds, absolutives, or whatever they are called) in strings before the predication-closing finite verb form or other predication-closer. It is certainly almost pan-Indic, although it may also be almost universal in

Eurasia, except for Semitic and Indo-European exclusive of Indo-Aryan. It is notable that Jules Bloch (*L'indo-aryen du veda aux temps modernes*, p. 285) was able to point out that these non-finite verb forms are lacking in modern Indo-Aryan in the Gypsy dialects, which presumably have lost this trait, along with so much else of their Indo-Aryan heritage, because of long separation from the Indo-Aryan languages of India. Otherwise, the only modern Indo-Aryan languages that lack the trait, on the basis of such evidence as is available, are Pashai, Tirahi, and the Kohistan group. These abut on Iranian territory on the borders of Afghanistan, and in the Iranian family the trait did not develop.

It is clear, on the basis of the two examples just given, that for any pan-Indic isogloss whose exact geographical spread one wishes to determine, one must examine the Indo-Iranian linguistic frontier carefully. For the two traits already examined, the isoglossal boundary at this frontier is drawn in two different places, one considerably west of the other. In one instance, that of the retroflexes, Indic influence has pushed the line into Iranian territory; in the other, Iranian influence has presumably caused loss of an otherwise pan-Indic feature and pushed the line back into Indo-Aryan territory.[3] These two form an interesting example of the failure of isoglosses to bunch.

Another isogloss discussed in the earlier paper was concerned with classifiers or quantifiers. This time the isogloss line includes part of northern and northeastern India but fails to coincide with the eastern boundary of the Indian sub-continent, however this may be defined. In fact, though this isogloss is an excellent one to use in

---

[3]It should be noted that, with Morgenstierne, I take the "Dardic" languages of Grierson to be, not a separate sub-family within the Indo-Iranian family, but a part of the Indo-Aryan sub-family.

demonstrating that a trait is shared by parts of the major families of India, it fails lamentably in demonstrating that India is a linguistic area and may be interpreted as showing that there are linguistic traits that occur in common in India and the rest of Asia.

At the end of the earlier paper (p. 16) I suggested that in the matter of the classifiers there is connection between India and Southeast and East Asia. I did not develope the matter there, but the data and their interpretation are relevant here.

The classifier system is one in which all or nearly all nouns are divided into subclasses, membership in which is indicated by the use of classifier morphemes when the nouns are numerated.

The Magadhan languages of Modern Indo-Aryan, i. e. Bengali, Assamese, Oriya, and, in this matter, some dialects of Bihari, have a fairly richly developed classificatory system, with a half-dozen or more classes and classifiers. The morpheme order is noun + numeral + classifier, or numeral + classifier + noun. Marathi has one classifier only, used with the subclass of nouns denoting persons (*jaṇ*) when such nouns are numerated by numerals higher than four (and optionally for two to four). The order is numeral + classifier + noun. Hindi in many dialects has a similar system, as do probably other modern Indo-Aryan languages. I assumed that possibly the Marathi system is ultimately derived from the richer Magadhan one. Of the Dravidian languages, Kolami (Wardha dialect) has borrowed from Marathi, and Kui-Kuwi, Kuru*kh*, and Malto owe their systems to Magadhan languages. The Malto system is developed to a greater extent than that of the other Dravidian languages. It has borrowed a number of Magadhan classifier morphemes, but also uses others of Dravidian origin. The order for the higher numerals is numeral +

classifier + noun, but for one and two it is classifier + numeral + noun. Several of the Munda languages have systems which seem to be borrowed from Indo-Aryan; they may be ignored here.

The rich Magadhan system seems to be the origin of the others in India, i. e. it is the centre of diffusion. However, when we collect evidence of systems of classifiers elsewhere in Asia, the Indian systems look like a western outlier of an area whose centre is East and South-east Asia. The evidence is as follows.

I shall deal first with languages of India outside the three major families. Burushaski does not have a classifier system, nor do the Iranian languages to the west of India. Khasi in Assam has a system of two classifiers: *ŋut* for persons, *tllii* for non-persons, the commonest syntactic construction being: numeral + classifier + plural indicator + noun (e. g. *aar ŋut kii briw* 'two men', *aar tllii kii dieŋ* 'two trees').[4] The Tibeto-Burman languages in the mountains to the north and east of the Indian plains are on the whole badly described, but their well-developed classifier systems could not easily be entirely missed and the *LSI*, vol. III, gives a fair amount of information about some of them, though there are still numerous gaps in the record. It is clear that some languages of the Himalayan group (e g. Newari) have classifiers, with the order numeral + classifier + noun, and that to the east all the groups, Bodo, Naga, and Kuki-Chin, include many languages so characterized, with order noun + classifier + numeral, or, in Bodo at least optionally, classifier + numeral + noun, for which cf. Malto above. In the islands of the Bay of Bengal, Andamanese apparently has no classifiers, but Nicobarese does (numeral + classifier + noun).[5]

---

[4]Information from Khasi informants and Dr. Lili Rabel.

[5]F. A. de Röepstorff, *Vocabulary of Dialects Spoken in the Nicobar and Andaman Islands*, 112 (1875). The details in the account are not very clear.

To leave India — Tibetan does not have a classifier system, either in the literary language or in the colloquials. Burmese, however, does, the order being noun + numeral + classifier.[6] Similarly for Thai in all its varieties (Siamese, Laos, Shan, Karen),[7] Khmer,[8] and Cham.[9] Vietnamese has numeral + classifier + noun [10] Curiously, archaic Chinese has no classifiers, classical Chinese has noun + numeral + classifier, colloquial Chinese has numeral + classifier + noun.[11] Korean has noun + numeral + classifier.[12] Japanese has two orders, numeral + classifier +

---

[6]William Cornyn, *Outline of Burmese Grammar*, §§ 141–8 (1944).

[7]Mary R. Hass, 'The use of numeral classifiers in Thai,' *Language* 18·201–5 (1942). For 'one' the order may also be noun + classifier + numeral.

[8]Georges Maspero, *Grammaire de la langue khmère* (*cambodgien*), 294–311 (1915).

[9]Étienne Aymonier et Antoine Cabatan, *Dictionnaire cam-francais* xxv *et passim* (1906).

[10]M. B. Emeneau, *Studies in Vietnamese* (*Annamese*) *Grammar*, chapter III (1951).

[11]R. A. D. Forrest, *The Chinese Language*, 112, 156 f. (1948); Edward H. Schafer, Jr., 'Noun classifiers in classical Chinese,' *Language* 24.408–13 (1948); Yuen Ren Chao, *Mandarin Primer*, 45, § 22 (1948; 'auxiliary nouns'); id., *Cantonese Primer*, 83 f. (1947).

[12]Fred Lukoff, *Spoken Korean*, 218 ff., 715 f., *et passim* (1945–7).

noun, and noun + numeral + classifier, the difference being apparently between definite and indefinite reference respectively, though the structure is complicated by the use of certain particles with the two constructions.[13] I hope that I have not missed the essence of the matter. Ainu has a scanty system with order noun + numeral + classifier.[14] In the Palaeo-Siberian group of languages, only Gilyak seems to have what we are looking for: there are some 25 classes of noun (I do not know what the morpheme order is).[15]

So much for mainland Asia and Japan. Altaic-Mongolian and Semitic do not have the phenomenon in question, nor does Indo-European in general.

The island world covered by Malayo-Polynesian can be surveyed in only the sketchiest way; only the specialist could with any ease go through the voluminous literature looking for this one small feature. *Les langues du monde* (1952), p. 659, assures us that all the Indonesian languages, including Malagasy, have classifiers and that the order is noun + numeral + classifier. Malay certainly agrees with this statement, but it also has, according to Winstedt, the order numeral + classifier + noun when the numeral is one, and a reversal in either case is said to involve emphasis on the numeral.[16] It is probable more-

---

[13]Bernard Bloch and Eleanor Harz Jordan, *Spoken Japanese*, 226 ff., 241 ff., *et passim* (1945); Samuel E Martin, *Essential Japanese*, 174 ff. (1954).

[14]John Batchelor, *A Grammar of the Ainu Language*, in Basil Hall Chamberlain, *The Language, Mythology, and Geographical Nomenclature of Japan viewed in the light of Aino Studies*, 95 f. (1887).

[15]Communication from Robert Austerlitz; also R. Jakobson, *American Anthropologist* 44·616 (1942).

over that it is only western Indonesian that has the feature and that it fades out in the sub-Philippines area from Borneo east. The best-known languages of the Philippines, e.g. Tagalog and Ilocano, clearly have no classifiers; they are not found in Formosan. The Micronesian languages have classifiers, e. g. Trukese, with order numeral + classifier + noun.[17] The Melanesian and the Polynesian languages do not, with some minor exceptions, such as a few of the Melanesian languages on the shores of Torres Straits in Southern New Guinea (Kiriwina—classifier+numeral+noun; Motu, Hula, Keapara—noun +classifier+numeral), and possibly Samoan.[18]

In this survey we have partially plotted on the map the distribution of a structural feature that possesses a high degree of specificity. The distribution has covered an immense area and a number of genetically unrelated language families Yet it has been limited to this one area of East and Southeast Asia and the contiguous island world—for, even if the feature is found elsewhere, e g. in Mayan, in Yurok, and elsewhere in the New World, it is not found elsewhere in the Old World to the best of my knowledge. There can be no question of reconstructing for all the families involved one parent language possessing the feature in question. The alternative explanations are an independent invention for each

---

[16] William Marsden, *A Grammar of the Malayan Language*, 30 f., 104 f. (1812); R. O. Winstedt, *Malay Grammar*, 129–31 (1913); Otto Dempwolff, *Einführung in die malaiische Sprache*, 48 f. (1941); Isidore Dyen, *Spoken Malay*, 114 f., 120 f. (1945).

[17] S. H. Elbert, *Trukese-English and English-Trukese Dictionary*, 22–5 (1947).

[18] *Reports of the Cambridge Anthropological Expedition to Torres Straits*. Vol. III. Linguistics by Sidney H. Ray, 475 (1907). On Samoan, information from Dr. Jesse Sawyer.

family concerned, or borrowing, i.e. diffusion, from family to family. Or, unfortunately for those who like cleancut solutions, there may well be an historical mixture of the two – since, considering that there is at least one independent invention of the feature in the New World, and very probably more, independent invention cannot be ruled out here either. We must proceed cautiously. And I find it the more difficult to proceed at all, since I cannot control the history of the literary languages of long standing, most importantly Chinese and Japanese. Perhaps it will be best not to try to set forth an over-all solution; it may be useful however to point to some not very profound, but nevertheless seemingly significant possibilities.

One of the obvious points of attack on the problem is an examination of the order of the morphemes. We find, very strikingly, a great central bloc of languages with the order noun + numeral + classifier — classical Chinese. Japanese (with indefinite reference, which is probably the most common construction), Ainu, Korean, and to the south Burmese, the Thai languages, Khmer, Cham, and the western Indonesian languages. Whether there is here a simple relation of diffusion from one source, as one would like to find, is not an easy question. The philological questions that must be asked about the history of and the relationships between Chinese, Korean, Ainu, and Japanese, cannot be answered by me, nor do I know whether they have yet been answered by the specialists in these languages. What is the relationship between the Thai languages and Chinese? What are the relationships, genetic and diffusional, between the languages of Southeast Asia and between those of Southeast Asia and those of Indonesia? What would reconstruction within Indonesian and further within Malayo-Polynesian have to say about the classifiers? They are probably not even proto-Indonesian, much less proto-Malayo-Polynesian, if we can trust the prima facie evidence of the distributions within

this family. But until the specialists supply information, the matter cannot go forward.

But we must attempt to connect the languages of India with those further east. The Indo-Aryan languages with which we started are not now contiguous with the central bloc; they are separated from it by the Assamese Tibeto-Burman languages with their order noun + classifier + numeral, even though one may note that this is one of the possible orders for the numeral 'one' in the Thai languages. There was, however, historically contact between the Magadhan area and a Thai language, Ahom,[12] from about A.D. 1228 until the extinction of this language in the 18th and the early part of the 19th century. Political dominance of the eastern part of the valley of Assam by Thai-speaking Ahoms for several centuries hardly seems enough to explain why all the Magadhan languages should have a classifier system with morpheme order like that of the Thai languages, and yet this is clearly a link that might have allowed diffusion from the Burmese Thai-Khmer area into India. Should we perhaps guess that the Ahom were only the last in a series of Thai-speaking communities who came into contact (unrecorded or unrecognized in our bad historical records) with speakers of Eastern Indo-Aryan in a proto–Magadhan stage? Yet what has been inferred about the linguistic prehistory of the present Magadhan area from the small islands of non-Indo-Aryan speech still found there, is that this was a Bodo, i.e. Tibeto-Burman, area with

---

[19]G. A. Grierson, *Linguistic Survey of India.* Vol. II. Mōn-Khmēr and Siamese-Chinese-Families, 59–140, esp. 94 (1904). The morpheme order is as in Siamese, even to the two optional orders when the numeral is 'one'. For the history involved, see also Suniti Kumar Chatterji, *Journal of the Asiatic Society of Bengal*, Letters, 22·147–53 (1956).

order noun + classifier + numeral, and not a Thai area, when Magadhan expanded into it.[20] However this may be, the Ahom or any other hypothetical Thai speakers will only account for the construction with order noun + numeral + classifier. For the order numeral + classifier + noun, which is found in Marathi, Hindi, and optionally in Magadhan, we must look to some other origin than a Thai one. Is this independent invention? I am loathe to accept it *tout simple*. Perhaps it is a case of Kroeber's 'stimulus diffusion' working westward from the Magadhan area, the old Indo-Aryan order numeral + noun (without a classifier in the construction) having been too strongly entrenched in the system to allow anything but the infixing of a classifier. This partially independent development may then have worked back again into the Magadhan area to provide the option found there. But I confess that I should like more good descriptive material from all over North India before I made up my mind too firmly about these historical hypotheses.

Other features of the classifier construction could be examined. E.g., which of the languages use classifiers in construction also with demonstrative words, as do colloquial Chinese, Thai and Vietnamese? What is the order of the morphemes in this type of construction, and so on? This information is even more difficult to find than that concerning the numerals, and I have not attempted to collect it.

Re-examination, then, of several isoglosses that delimit linguistic areas either within India, or roughly pan-Indic, or in which part of India is included, has probably yielded the differentiating feature that distinguishes a multifamilial area (or even a multilingual area) from a multidialectal

[20] We must not of course forget that one of the Malto orders, classifier + numeral + noun, is also found in Bodo.

area. In the first case there is a tendency for each isogloss to establish a different area. This is, to be sure, true also in dialect studies and leads to such familiar features of our textbooks as the Rhenish 'fan' (and undoubtedly many others). But there is in dialect studies a clear bundling; a familiar example is that between Low and High German, of which the Rhenish fan is the frayed end. In linguistic area studies it is doubtful if there will ever emerge isogloss-bundles. I should not however stop without guessing that if a linguistic area *is* found that is clearly marked off by an isogloss-bundle, it will be India.

II

In the previous lecture two examples have illustrated our concern with the Indo-Iranian frontier area and the drawing of isoglossal lines which exclude Iranian languages from the Indic area. It has become clear in this operation that on this frontier we are dealing with an area of bilingualism (or multilingualism)[21] in which it is not pre-determined that Indic influences should be prepotent; for one trait there was such prepotency, for the other not. The structural feature that is the chief subject of the present paper reverses the direction of interest. A feature of structure which must surely be called 'intimate' in Sapir's sense, is almost pan-Iranian, and is found in the Dravidian Brahui and also in some of the Indo-Aryan frontier languages. The problem of the paper is first to demonstrate and analyse the synchronic situation, and then to make an assault on the historical problems involved.

---

[21]This statement is certainly nothing new. All who have written of this area, and one need name especially only Grierson and Morgenstierne, have both provided the evidence and made the general statement. The application here to the broad question of inter-influences is new in detail.

This trait is the use of what is most often called in the descriptions 'pronominal suffixes.' There are few well-structured accounts of the trait; in fact, for most of the languages involved it is possible only to recognize that it is there, but quite impossible to find accurate statements about it. The name used, 'pronominal suffixes,' is misleading. What is meant is not in general pronominal inflexions in the verb system, but another set of pronominal forms which are fairly easily distinguishable from the former by several criteria. We shall start by an examination first of the Iranian languages. Nearly all the modern members of the family (certainly at least in the great southern bloc) of which we have intelligible accounts have the trait. Ossetic seems not to have much that is equatable with what we are interested in, and will be left out completely in what follows. We shall be unable to examine all the languages; statements for some of them are so scanty that even the small amount of detail that we need to recognize and classify our trait is unavailable. Examination of a few of the best-described languages will be enough.

The pronominal suffixes in which we are interested show some or all of the following features. (1) In most of the languages past tenses of transitive verbs have a set of pronominal references to the subject which is different from that (or those) of all other verb forms. Standard Persian, whether literary or colloquial, seems to be aberrant in not showing such a difference; some even of the Persian dialects have it. The past transitive pronominal references are those we call pronominal suffixes. I must at times go into the question of whether the transitive past tenses are structurally passives; they are descriptively so in some of the languages, though not in others, and are

surely historically so in all of them.[22] The point is in part relevant to our purpose. (2) These pronominal suffixes often occur in a verb form as a second pronominal reference, usually to the object. (3) Besides their occurrence with verb forms, they often occur as possessives with nouns, and in construction with prepositions. (4) Often a pronominal suffix appears in phonological immediate constituency with a form other than that with which it is in immediate constituency morphologico-syntactically; this is true for the immediate constituencies of (1), (2), and (3). This feature applies in (1) to the subject reference of past transitive forms; this pronominal reference may appear, e.g., earlier in the sentence, and the verb form is then without pronominal suffix. This feature accounts for the frequent reference to the pronominal suffixes as 'enclitics.' (5) Often one of these suffixes is repeated in the sentence, or repeats a noun or an independent pronominal form. (6) In some of the languages some of the verbal paradigms, especially some of the preterites (aspect-modal-tenses), consist of what seem to be phonologically unified words made up of prefix (aspectual or modal) + pronominal suffix forms + past stem (± further elements). Historically these are structurally the same as prior word (or morpheme) + pronominal suffix, followed by the past stem.

For modern Persian the account by Lazard seems full. He treats the pronominal suffixes (pp. 101–9) as 'enclitic personal suffixes', distinguishing them from the personal

---

[22]See the excellent summary article by Wilhelm Geiger, 'Die Passivconstruction des Präteritums transitiver Verba im Iranischen,' *Festgruss an Rudolf von Roth*, 1–5 (Stuttgart, 1893). Cf. also Constantin Regamey, 'A propos de la "construction ergative" en indo-aryen moderne,' *Sprachgeschichte und Wortbedeutung, Festschrift Albert Debrunner*, 363-81 (Bern, 1954).

pronouns and the personal inflexions ('désinences personelles'). In the colloquial the forms are:

	Sg.	Pl,
1	C] *am*, V] ***m***	C] *emun*, V] ***mun***
2	C] *et*, V] *t*	C] *etun*, V] *tun*
3	C] *eś*, V] *ś*	C] *eśun*, V] *śun*

In the literary language and in formal speech the post-consontal forms usually occur after both consonants and vowels; the 2nd and 3rd sg. are *-at* and *-aś*; the vowel of the final syllable in the plural forms is *â* (the back *a*) These suffixes are much less frequent in the literary language than in the colloquial. Feature (1) is not found in modern Persian, in which the subject of the preterite, transitive as well as intransitive, is always marked by personal inflexions (not by pronominal suffixes) attached to the past stem. Older Persian texts, however, show this feature.[28] Features 2–5 are all found in modern Persian: (2) *did-am-et* 'I have seen you'; *bogu-ś biâd tu* 'tell him to come in' (*bogu-ś* tell-him); (3) *dast-am* 'my hand'; *be-m* 'to me'; (4) *ce-t śode* 'what has happened to you'? (*ce-t* what-to you); *dust-et dâram* or *dust dâram-et* 'I love you' (love-of you [] [] I have;

---

[28]Cf. the account by Paul Horn in *Grundriss der iranischen Philologie* 1·2·118–9. For the rest, this account is much shorter than Lazard's and shows some difference of detail; I do not pretend to be able to reconcile the two accounts. I note also the review of Lazard by Carleton T. Hodge in *Language* 34·111–21 (1958); he gives a further morphological analysis of the pronouns, which it is not necessary to reproduce in this connection.

love [] [] I have-of you); [24] (5) *to-ro* [= *to-râ*] *duss-et* [= *dust-et*] *dâram* 'I love you' (*-et* repeats *to-*). Feature (6) is not found, though some of the dialects have it; e. g. Lambton (p. 7) gives for Meime' *ī bɛmhɔru·t* 'I sold', *bɛthɔru·t*, *bɛśhɔru·t*, etc., the preterite of a transitive verb, and *amca* 'I was doing', *atca*, *aśca*, etc., the imperfect of a transitive verb.

Another language for which this feature is well described is Kurdish. McCarus describes the East Kurdish dialect of Sulaimaniya (taken as standard in Iraq). An article by C. J. Edmonds in *BSOAS* 17.490–502 (1955) describes 'Prepositions and personal affixes in Southern Kurdish'; the dialect seems to be the same as that described by McCarus. McCarus' account is the fuller and the more exactly stated. The pertinent forms, called by him 'pronominal suffixes' and by Edmonds 'pronominal affixes', are:

	Sg.	Pl.
1	C] *ɪm*, V] *m*	*-man*
2	C] *i*(*t*), V] *t*	*-tan*
3	C] *i*, V] *y*	*-yan*

(*i* is high, front, unrounded. *ɪ* is high, back, unrounded; in 2sg. (*t*) 'may or may not be pronounced'). McCarus defines the forms by saying that they 'may occur with any form class and ... do not convert the stem word to a different form class'. They show all six features given above. However, Kurdish shows considerably more complication than does Persian in the interplay between the pronominal suffixes and the other pronominal forms called

[24] Though standard Persian does not show feature (4) applied to (1), i.e. the pronominal subject reference of a past transitive appearing earlier in the sentence, many Persian dialects do; cf. some of the material given by Lambton, esp, pp. 7, 45,

verbal suffixes in their occurrence with transitive and intransitive verbs, and with the various tenses. According to feature (1), the pronominal suffixes show the subject with the past tenses of transitive verbs. All other verb forms are constructed with the verbal suffixes:

	Sg.	Pl,
1	C] *ịm*, V]*m*	C] *in*, V] *yn*
2	C] *i*(*t*), V] *y*(*t*)	C] *ịn*, V] *n*
3	*-e*(*t*), *-a*(*t*)	C] *ịn*, V] *n*

The 3sg. of an intransitive verb in the past tenses has a zero suffix. The usual shift of position of the pronominal suffix, which is feature (4), is seen in such a sentence as *márek-ịm küšt* 'I killed a snake' (*ma'rek-ịm* a snake-I). McCarus' section on the shifts of suffixes with prepositions is a good example of the complications that are found (pp. 105–6). Shifts occur when the verb is transitive (or, in the present tense, also the copula). When the tense is past, the subject suffix of the verb is replaced by a pronominal suffix added to a prior word in the sentence, and the pronominal suffix with the preposition is replaced by a verbal suffix with the verb. If prior to the preposition there is a word which is eligible to have a pronominal suffix (there is a rule determining eligibility), the preposition remains without any attached pronoun form; if the preposition itself is the first word of the clause, it has the subject pronoun. Exx. *cák-i ləgəl kịrd-ịm* 'he did well by me' (good-he [] [with [] [] did-me); *pé-yan baxś-im* 'they presented [it] to me' (to-they [] [] presented-me). When the verb is in the present tense, the pronominal suffix with the preposition is shifted to a prior word; e.g. *cirókek-ịt bo ʔəlé-m* 'I'll tell you a story' (story-you [] [] for [] [] say-I). Edmonds (p. 502) records a similar usage for the possessive suffix with a noun; e.g. *kaghez-eke-yan drhiyw-im* 'they have torn my letter' (the letter they [] [] have torn-my). Feature (5), repetition of

reference, is seen in a sentence given by McCarus (p. 104): *pyawəkán dəryan kịrd* 'the men removed it' (*dər-yan* out–they). Feature (6) is seen in the imperfect of transitive verbs (e. g. *ʔəmxwárd* 'I was eating,' *ʔətxwárd*, *ʔəyxwárd*, etc.), in the preterite subjunctive of transitive verbs (e.g. *bịmxwárdayə*, *bịtxwárdayə*, *bixwárdayə*, etc.), and in other paradigms.

Of the eastern languages perhaps the best description is that of Parachi (just north of Kabul) by Morgenstierne (1929). The account is not as detailed as those for Kurdish. The statements given are, however, sufficient to show that, apart from the forms themselves, the structural features are more like those of Kurdish than of modern Persian. The forms in question are:

	Sg.	Pl
1	C] *um*, *əm*, V] m	C] *an*, V] *n*
2	-*au*[V, -*a* [C	-*ōu*, -*ō*
2	-*ē*	-*an*, -*ē*

The first five features are found in the system. Exx.: (1) contrast *ȷan-em* 'I kill' and *ȷö-m* 'I killed,' *xar-em* 'I eat' and *xūṛ-um* 'I ate'; (2) *mēran-um te* 'they will kill me' (*te* is of uncertain meaning); *mērem-a te* 'I shall kill thee'; (3) *biyā-m* 'my brother'; (4) *derâk-um dâ* 'give me a grape'; *xöm-um dhöṛ* 'I dreamt' (dream-I [] [] saw); *ce-um qasam xūṛö* 'that (conj.) I have sworn' (that-I [] [] oath [] [] ate); *tū kun-əm kitab dâ* 'I gave thee a book' (thee [] [] to-I [] [] book [] [] gave); *aze-m dhöṛ-a* 'I saw thee yesterday' (yesterday-I [] [] saw-thee); (5) *manân nhâmuṛ-um chī* 'I have forgotten' (my [] [] forgetting-my [] [] is); *ma bâw-aw-an ânt-an* 'we (or they) brought thy father' ([accusative marker] [] [] father-thy-we (or they) [] [] brought-we (or they)). There are no indications of any forms illustrating feature (6).

Ormuri, south of Kabul, described by Morgenstierne (1929), has features 1–5. One form is to be noted, 2sg. C]*at*, V]*t*, for its retention of *t*, which Parachi does not have.

Of the Ghalchah languages in the northeast corner of the Iranian territory, the Hindu-kush and Pamir region, the material is on the whole still scanty. The *LSI* X gives some information. For Shigni (p. 467) it makes it clear only that features (1) and (4) are found. The *LSI* and Grierson (1920) guarantee features (1), (4), and (5) for Ishkashmi; Grierson speaks of a " wider use " in Zebaki, adding feature (2) in an example. Wakhi, according to Morgenstierne (1938), pp. 492, 505–6, has features (1), (4), and (5). According to D. L. R. Lorimer, *BSOS* 3·485–6, and the 1958 publication, both intransitives and transitives have the same pronominal suffixes. A few of the sets of forms are given in table 2. It is necessary only to point out that in several of the languages the 1pl. has *–n*, the 2sg. *–t*, and the 2pl. *–v* or *–f*.

Of the eastern Iranian languages, Pashto and Balochi on the Indo-Iranian border, with their comparatively large numbers of speakers, are of considerable interest.

Pashto has been described often enough. The *LSI* account has, curiously, nothing of interest in this matter except a list of the pronominal suffixes. The most recent description, that by Penzl (Kandahar dialect), gives probably an exact account of the matter. As in most of the older accounts,[25] the forms are given by Penzl as enclitic particles rather than as suffixes. He takes account (p. 37, § 42·4b) of the phenomena of close juncture and lack of stress which in certain types of description (e.g. Nida's) would ensure the labelling of these particles as

---

[25]E.g. Wilhelm Geiger in *Grundriss der iranischen Philologie* 1·2·217; D. L. R. Lorimer, *A Syntax of Colloquial Pashtu.*

enclitics. Unfortunately, he fails to carry the stress and junctural statements throughout the grammar, so that it is impossible to quote the pronominal forms as anything but words. The forms in Kandahar are :

	Sg.	Pl.
1	*mi*	*mu*
2	*di*	*mu*
3	(*y*) *ee*	

Eastern dialect forms are: *me*, *dee*, (*y*) *ee*, *moo* or *um*. Pashto does not have feature (1), i. e. the past tenses of transitive verbs have essentially the same pronominal inflexions as all other paradigms. It is clear, however, that the past transitive is a passive. Feature (5) is not found, i. e. the repetition of pronominal references involving the forms in question. Feature (3) is clear ; e. g. *koor di* 'your house,' which according to §42·4b is *ko'ordi* (close juncture) or *ko'or-di* (hyphen juncture). Feature (2) appears when the imperfective present and a pronominal particle are the only forms in the utterance (p. 134, §100·5c) ; e. g. *winem di* 'I see you,' in which there can be presumed to be close (or hyphen) juncture. In all other constructions, however, the particles precede the verb form somewhere in the utterance, consequently show features (4) and (6); e. g. *oobe di raawrree* 'water was brought by you' (*di* by you ; p. 133, §100·2b ; probably *o'obedi*) ; *neewini* (close-junctural resultant of *ne yee wini*) 'he does not see him' (not [] [] him [] [] he-sees ; p. 87, §78·2).

Descriptions of Balochi are less than satisfactory. I rely on *LSI* X.344-5, Wilhelm Geiger, *Grundriss der iranischen Philologie* I. 2·240, and Gilbertson, *The Balochi Language*, pp, 70–4, 117–9. There are some

points on which they differ, and no reconciliation seems possible. The forms in *LSI* are:

	Western	eastern
1sg. pl.	*–u*^{n}	*–ā*^{n}, *ū*^{n}
2sg. pl.	*–it*	?
3sg.	*–ē, i* pl *–ish*	sg. *–ī* pl. *–ish*, *–ā*^{n}

Gilbertson gives forms which are apparently western:

1sg.	*úṅ* [presumably *LSI*'s *u*^{n} or *ū*^{n}]	
2sg. pl.	*e*	
3sg.	*ì*	pl. *ish*

Geiger, depending on an earlier source (Mockler or Dames) has:

1	*–un*	
2	*–it*	
3sg.	*–ē*	pl. *–iś*

It is not quite certain what the state of affairs is as regards feature (1); certainly the past tenses of transitive verbs seem to be passive in construction. Gilbertson's ex. *tu'fak gipto a'rth-uṅ* 'I have brought the rifles' (rifles [] [] having-taken [] [] I-brought [or, have-been-brought-by-me]), looks like (1), but there can be no certainty. Feature (2) is clear; e. g. (Geiger) *man kusān-ē* 'I will kill him'; (Gilbertson) *bozhīth-ish* 'let him untie them'; *zar de–a'n–ish*, I will give them money.' So is (4); e. g. (Geiger) *zar–ē bidai* 'give him money' (*zar–ē* money–to him); *LSI* (p. 345) quotes an example from Mockler. Features (5) and (6) seem not to occur. According to Gilbertson, (3) does not occur ("found with verbs only"), but Geiger and *LSI* give clear examples and make explicit statements that the suffixes are found with nouns; e. g. *LSI mālai* (< *mālā–ē*) 'his cattle'; *girōx––ī* 'a purchaser of it' (*ruskat–ē* '[give] leave to him' is more likely an example of 4).

On the Indo-Aryan side of the linguistic frontier phenomena similar to those we have been examining on the Iranian side are restricted without exception[26] to languages that lie on or near the frontier itsef. Sindhi and Lahnda have them. So also have most of Grierson's Dardic group, including Kashmiri and Shina of the Dard group (but not Kohistani) and all the Kafir languages but Wasi-veri (otherwise called Prasun); Khowar, the farthest northwest of the Dardic group, does not have the phenomena in question. It is to be noted, however, that the negatives in the Dardic group are not to be taken as certain; the descriptive material at hand is very scanty.

In Sindhi, for which the account in *LSI* VIII.1.42–5, 65, 67, 78–91, 93-5, is full, features (1), (2), and (3) are found. Two sets of pronominal suffixes occur:

	A		B	
	sg.	pl.	sg.	pl.
1	$-s^{e}$	$-s\bar{i}^{n}$, $s\bar{u}^{n}$	$-m^{e}$	$-\bar{u}^{n}$, $-s\bar{i}^{n}$, $-s\bar{u}^{n}$
2	$-\bar{e}^{n}$(fem. –an)	$-u$	$-^{e}$	$-w^{a}$
3	—	—	(a) $-s^{e}$	$-n^{e}$, $-n^{a}$
			(b) $-\bar{a}\bar{i}^{n}$	$-\bar{a}\bar{u}^{n}$

1pl. $-s\bar{i}^{n}$ and $-s\bar{u}^{n}$ are said to be in free variation; they occur (generally) after vowels and $-\bar{u}^{n}$ after consonants. Feature (1): Set A is found as subject in certain of the tenses based on participles, viz. the future, the past of intransitive verbs, and the past of the passive; e.g. *handus*^{e} 'I (masc.) shall go,' *halius*^{e} 'I (masc.) went,' *mārius*^{e} 'I was struck.' Set B(b) is found in the past of transitive verbs, which is a passive construction, the suffixes indicating the agent; e.g. *māriāī*^{n} 'he or she struck

---

[26] The possessive suffixes found with relationship terms in Assamese are historically of such a different origin from those of the northwest frontier that they may be ignored here.

(him, her, or them),' *mārium*^{e} 'I struck him.' (2) Set B(a) is found as a second (or even third) pronominal reference in verb forms, equivalent to any oblique case except the agent (of the past of transitives); e.g. *jhalī*^{n}*–s*^{e}, *hē mārē–m*^{e} *thō* 'seize him (*–s*^{e}), he is beating me (*–m*^{e})'; *ḍḍakē–w*^{a} 'he trembles at you'; *ḍḍināī*^{n}*–m*^{e} 'he or she gave [it or them] to me'; *mānī khādhī–mā*^{n}*–*^{e} 'I have eaten your bread' (bread [] [] has–been–eaten–by–me–your; *-m*^{e} is replaced by *–mā*^{n} before another suffix). (3) Set B(a) occurs also with nouns (but lpl. never occurs), and with forms that otherwise occur as postpositions after the oblique case of a noun; e.g. *piu–m*^{e} 'my father,' *dhaṇya–n*^{e} 'their master,' *nēṇu–*^{e} 'your eye'; *saṇu–m*^{e} 'with me,' *kaṇe–ū*^{n} 'to us.'

The *LSI* account for Lahnda is in VIII.1.236, 259-62, 267, 269–71. It is not as full as that given for Sindhi; the system however seems in general parallel to that of Sindhi, with features (1), (2) and (3). Exx.: *ghara–m* 'my house,' *nā*^{n}*–ī* 'your name'; *mä*^{n} *māreā–s* 'I struck him' (by–me [] [] he–was–struck); *mänū*^{n} *māreā–s* 'he struck me' (to–me [] [] it–was–struck–by–him); *rupaiā dittē–mu–s* 'I gave him a rupee' (rupee [] [] was–given–by–me–to–him). The two sets of endings given are:

	A		B	
	sg.	pl.	sg.	pl.
1	–s	–s	*–m*	*–sē*, *–ahsē*
2	*–ē*^{n}	—	*–ī*, *–ū*^{n}, *-ō*, *ō*^{n}	*–nē*, etc.
3	—	—	*–s*	*–nē*

The status of set A is not clearly parallel to that of A in Sindhi. It is not clear whether the pronominal suffixes occur with postpositions. Notes are given on the occurence of the suffixes in the northeastern and northwestern dialects of Lahnda, including (p. 545) the statement that the 2sg. nominative suffix is *-j*.

The account for Kashmiri in *LSI* VIII.2.282, 290–304, 312–4, is full; it is based on the careful statements of native Kashmiri grammarians. The pronominal suffixes do not occur with nouns, nor apparently with postpositions, i.e. feature (3) is not found. The features that are found are (1), (2) and also (5). The forms are :

	A		B		
	sg.	pl.	sg.		pl.
			acc.-agent	gen.-dat.	
1	*–s*	—	*–m*	*–m*	—
2	*–kh*	*–wa*	*–th*	*–y*	*–wa*
3	—	—	*–n*	*–s*	*–kh*

When either of the suffixes *–kh* is followed by another pronominal suffix, it is replaced by *–h–*. After a consonant, the suffixes are *–as*, *–akh*, etc. Suffixes of set A are found as subject in the tenses derived from the past participles of intransitive verbs; e g. *wuɒu–s* 'I (masc.) was burnt inwardly (proximate past)'. The tenses derived from the past participles of transitive verbs are passive constructions; e.g. *wucchi–m* 'I saw them' (they were seen by me). With the participles are found an agent suffix from set B, optionally followed by a subject suffix from set A; the missing 3rd persons of set A are replaced by the acc.–agent forms of set B, *–n* and *–kh*; e.g. *wucchᵘ–h–as* 'they saw me' (seen–by–them–I); *wucchᵘ–th–akh* 'thou sawest them' (seen–by–thee–they); *wucchᵘ–n–as–ty–ā* 'did he see me also ?' (seen–by–him–I–also–?). Set B is otherwise found expressing oblique cases added to verb forms, according to feature (2); e.g. *di–m* 'give to me'; *wuphyō–h–am* 'thou flewest for me' (flew–thou–for–me); *wucchyū–n* 'you will see him'. One complication in statement is that the gen.–dat. suffixes occur instead of the acc.–agent suffixes after a 3rd person; e.g. *wucchi–y* 'he will see thee,' *wucche–s* 'he will see him.' Feature (5), the repetition of pronominal references, occurs with some

complication. A pronoun or noun of the 1st or 3rd persons preceding the verb is optionally repeated by a pronominal suffix; e.g. *me wucch*ᵘ or *wucchu-m* or *me wucchu-m* 'I saw him' ([by-me] he-was-seen-[by-me]). When a 2nd person pronoun precedes the verb, it is obligatory that a 2nd person pronominal suffix appear; e.g. *wucchu-th* or *tse wucchu-th* 'thou sawest him' ([by-thee] he-was-seen-[by-thee]); *wana-y* or *tse̱ wana-y* 'I shall say to thee' ([to-thee] I-shall-say-[to-thee]).

For Shina, which is placed close to Kashmiri in the Dard sub-group of Dardic, the accounts are not good on this trait. The *LSI* VIII.2.217-8 makes it clear for one dialect that features (1) and (2) are found; it is fairly clear that (3) is missing, just as in Kashmiri. Bailey, p. 218, gives very scanty information on another dialect; D. L. R. Lorimer, *BSOS* 3 467-93 ('The forms and nature of the transitive verb in Shina') has nothing relevant.[27]

On the Kafir group, the information is sketchy.

Shumashti, a dialect of Gawar-bati, is described by Morgenstierne in *NTS* 13·239-81. The pronominal suffixes are shown (pp. 254-5) as having probably feature (1), and certainly features (2), (3), and (5). The forms are:—

	Sg.	Pl.
1	*-m*	*-ani/e*
2	*-u*	*-i*
3	*-s*	*-at*

Varying vowels precede the consonants of the suffixes, depending apparently on classes of the preceding forms. Exx.:— *nām-am* 'my name,' *āŋuṛ-em* 'my finger,' *id-əm* 'my heart,' *zū-im* 'my daughter'; *gäiw-am-a* 'didst thou see me?' (*-am-* me); *piēniāw-ane* 'you know us' *mono bā-m* 'my father' (my [] [] father-my); *mō tharalas--em* 'thou seest me' (me [] [] thou-seest-me).

[27]For Kohistani, also included by Grierson in the Dard sub-group, nothing is recorded either in the *LSI* VIII·2, or in Grierson's *Torwali*.

Of the Kalasha–Pashai languages (if that is the correct grouping), Pashai certainly has features (1), (2), (3) and (5), with no information as to whether the last occurs only with nouns or with verbs as well. However, we are still waiting for Morgenstierne's promised Pashai grammar. In his 1944 volume of texts, an inserted slip says that the grammar and vocabulary are in preparation; no details from the notes published so far in the text volume and in the 1926 *Report*, pp. 92–3, need be given.

Bashgali is described in *LSI* VIII.2 without any fulness. There are sufficient indications (pp. 36–7, 41) to show that features (1), (2), and (3) are found, and that feature (5) is seen in constructions with nouns. Morgenstierne prefers the name Kati for this and related dialects (1932, p. 63). He gives some Kati forms and constructions in his paper on Ashkun; the details differ somewhat from those of the *LSI*, but the system seems to be the same.

For Waigali (Wai–alā) sources are *LSI* VIII. 2 and Morgenstierne, *NTS* 17.146–324. The system of this trait is still however uncertain. Feature (3), i. e. use of possessive suffixes with nouns, occurs combined with (5). The *LSI* gives a few examples which point to use of these suffixes with verbs, i. e. feature (2); (1) is quite uncertain.

On Ashkun almost the only information is that given by Morgenstierne, *NTS* 2.192–289. He describes (pp. 211–2) possessive suffixes which exemplify features (3) and (5).

Finally, the Brahui language of the Dravidian family is surrounded by Balochi, i. e. Iranian, territory. It too has pronominal suffixes of the type that we have been describing. A description is found in Bray, 1.78–82. There are enclitic pronouns which occur suffixed to verbs and to nouns (features 2 and 3). Feature (1), found in both Iranian and Indo–Aryan and usually involving pronominal forms for transitive verbs in the past tense(s)

that are different from those of other verb forms, is, if not quite absent from Brahui. at least not related to the pronouns under discussion. Certainly the important feature (or origin) of the Indo-Iranian system, that the past of transitive verbs is a passive construction, is not found in Brahui. Features (4), (5), and (6) are all missing from Brahui.

Brahui, however, has its own features that are not found elsewhere. As a corollary of the lack of features (1) and (5), the enclitic pronouns have no nominative forms. They represent genitive, dative and accusative cases. The forms are :

	Sg.	Pl.
1	*-ka*	*-nan*
2	*-ne*	*-num*
3	*-ta(n)*	*-tā(n)*

The form *-tan* is an "occasional variant" of *-ta* ; *-tan* is given in Bray's vocabulary. The 1st and 2nd sg. forms are said to be "not infrequently used, especially in Jhalawan", i.e. the Brahui south of Kalat. The 1st and 2nd pl. forms are found "in some dialects", e.g. that of the Nicharis (who are Jhalawan), but are "inadmissible in the standard dialect", which is that of Kalat, to the north of the Jhalawan territory.

Exx. : *bāva-ta pārē* 'his father said'; *bāva pārē-ta* 'the father said to him'; *khalkuṭ-ta* 'I struck him'; *khalkus-ka* 'you struck me'; *khalkus-nan* 'you beat us' *pārē-num* 'we said to you (pl.)'; *karife-ne* 'he will make thee (so-and-so)'.

For any other oblique case, a base *i* which is declined with the ordinary case endings, has suffixed after the case-ending an enclitic of the 1sg., 2sg., or 3rd personal

pronouns, or a genitive of the independent 1st or 2nd personal pronouns, according to the following table:

		Sg.	Pl.
i + case-ending +	1	*-ka, -kanā*	*-nanā*
	2	*-ne, -nā*	*-numā*
	3	*-ta(n)*	*-tā(n)*

The genitive of the reflexive pronoun also is found in this construction: *i* + case-ending + *-tēnā*. Exx.: *masiṛ hinā itō-tā* 'the girl went with them'; *dudēngā iai-ta* 'he ran towards her'; *asiṭ iṭī-ṭā* 'one among them'; *nī ītō-tan hinak* 'you go with him'; *īki-numā ant hēsur* 'what did they bring for you (pl.)?'; *īis-kv harci ki khanis nā ē* 'all you may see in my possession (*īis-ka*) is yours'; *i dā giṛāte muccāi iki-tēnā halkunuṭ* 'I've purchased these things one and all foe myself (*iki-tēnā*).' Bray's account, excellent as it is in general, leaves some points unclear; one is the difference in usage between case-forms of the independent pronouns and those of the enclitics, especially when the latter are suffixed to the base *ī*.

We have now drawn an isogloss that encloses most of the Iranian languages, the Dardic languages, Lahnda, Sindhi, and Brahui, and have analysed the relevant features language by language. We may note that the eastern boundary of this isogloss is much further east than either of the two discussed earlier, and then proceed to attempt historical analysis.

Table I indicate which of the features of the analysis is found in each of the languages for which there are relatively full accounts. Contrastive distributions are very obvious. It is only Iranian languages that have all

six features, viz. Persian when taken as a whole, and Kurdish. That is, older Persian had feature (1) and modern standard Persian has now lost it: feature (6) is found in some Persian dialects, though not in modern standard Persian. These two languages are in the central and western part of the modern Iranian territory, The eastern languages for which there are sufficient data show fewer of the features. Parachi and Ormuri lack (6); Pashto lacks (1) and (5); Balochi, in which we are particularly interested because of its geographical relation with Brahui, is badly described and seems to lack (1), (5). and (6). All of the languages for which there is certain information have (2) and (3), and all have (4), i.e. pronominal suffixes which occur with nouns, which occur with verbs as second pronominal references, and which show failure of agreement between their immediate constituencies phonologically and morphologico-syntactically.

Nearly all the Indo-Aryan enclosed within the isogloss, for which there is clear information, have features (1), (2), and (3). Kashmiri and Shina are peculiar in that they lack (3), i.e., do not have the suffixes as possessives with nouns. Ashkun is peculiar in apparently having this feature alone. The Kafir languages are almost alone in having feature (5), i.e., repetition of pronominal reference; this in general seems to be found in these languages only in the possessive construction. Kashmiri is peculiar in having feature (5) in the verbal construction, and that in an unsymmetrical way, in that it is obligatory when the 2nd person is involved and optional for the other persons; this is undoubtedly a special development in this language alone. None of the Indo-Aryan languages has feature (6), and basically more significant, none has (4), i.e. freedom in immediate constituencies, which is so characteristic of Iranian. Finally, these Indo-Aryan languages are alone in the Indo-Aryan family in having any of the apparatus, except for (1), which is almost pan-Indo-Iranian.

Brahui, the Dravidian language, does not have (1), nor (4), (5), and (6). It agrees with the Iranian and Indo-Aryan languages examined in having (2) and (3), i. e. pronominal suffixes which appear both with nouns and verbs. It has its own special features, viz. lack of forms used in the nominative case, and a base inflected for case to which the pronominal forms are suffixed when the case is other than genitive, dative, or accusative. It is of importance to note that Brahui is the only Dravidian language that has any of this structure.

Considering that the structure we are concerned with is almost pan-Iranian and that the Iranian languages show greater complications in the trait than the Indo-Aryan languages or Brahui, we should be tempted to conclude that historically the trait originated in Iranian and spread thence to the Indo-Aryan languages on the Iranian frontier and to Brahui which is in Iranian territory. In other words, for this isogloss Iranian might be judged to be the prepotent element in the linguistic area. However, this is to jump at conclusions on the basis of distributions alone, and there are other data which in fact change the direction of our conclusions.

The basic assumption that we make in our historical analysis is the diffusionary assumption of anthropologists and linguistic scholars both. If a cultural trait (and this includes linguistic traits or structures) of any complexity of detail and/or organization is found in two or more contiguous communities, it is assumed to have only one place of origin and to have spread thence by diffusion to the contiguous communities. This is only the first and simplest case in question, and the statement must be qualified in various ways for special cases in which we are able to add more data. For example, if all of the communities can be proven to be genetically descended from one earlier community, inheritance only need be

invoked, unless there is evidence to the contrary, i. e. evidence for diffusion within this genetically related area. If some of the communities in a linguistic area are genetically related, we can, though we need not, assume inheritance for them and diffusion for the others which are not included in the genetically related group; the real problem here will be identification of the originating community, which need not be, nor be included in, the genetically related group of communities. It is the last case that we should assume here, though with a further complexity.

We have found a linguistic trait of some complexity of detail and organization in a large area, in which there are two large genetic groupings of languages, Iranian and Indo-Aryan, and one language genetically independent of both groups. We may assume origination of the trait in any one of the communities and diffusion thence to all the others. Origination in Brahui, however, is the assumption that has least probability, since the trait appears nowhere else in the Dravidian family, and there is in fact historical linguistic evidence that makes origination elsewhere in the linguistic area much more probable.

It may be well to finish consideration of Brahui at this point. One question that arises is how much of the total trait Brahui has borrowed. The diffusion, in fact, has been of structure only, and not of detail. All the enclitic forms involved have more or less cogent etymologies within Brahui itself and within Dravidian, and are not etymologically related to the Iranian or the Indo-Aryan forms. The 1st and 2nd person forms are all equatable with the oblique stems of the Brahui independent pronouns: 1sg. *kan-*: enclitic *-ka*; 1pl. *nan-*: enclitic *-nan*; 2sg. *ne-*: enclitic *-ne*; 2pl. *num-*:

enclitic *-num*. Dravidian etymologies are obvious for the last three (*A Dravidian Etymological Dictionary*, entries 3019, 3051, 3055) and probable, though problematical in detail, for the first (entry 4234). The 3rd person forms, Bray suggested, are derivable from the Dravidian reflexive pronoun (entry 2612); there is a possibility here, but no great probability; there is no probability at all that they are borrowed from either Iranian or Indo-Aryan. The base *ī-*, to the inflected forms of which the enclitics are suffixed, looks suspiciously similar to the Dravidian nearer demonstrative *i/ī* (entry 351); I have attempted to work out the details of this etymology elsewhere.[28]

In generalizing our description of the linguistic structures, we ended with features (2),(3) and (4) as pan-Iranian, (2)and (3) as found in nearly all the Indo-Aryan languages in question, and (2) and (3) as the only features in Brahui (apart from its own peculiar developments). Brahui's structure, therefore, is more nearly similar to the Indo-Aryan structures than to the Iranian. The specific point involved is its failure to show feature (4), which is a most important differentiating mark of Iranian as against Indo-Aryan. It seems that we should probably conclude on this ground that Brahui has borrowed the trait from Indo-Aryan. This may seen startling at first sight since Brahui is an enclave in Iranian territory, and more specifically in Balochi. On strictly linguistic grounds we may note that, however badly Balochi is described, feature (4) clearly belongs to it.

---

[28]Bray with his usual acumen suggested (I·15) that the use of the enclitics "points clearly to foreign influence," and (II·21) that the influence is "Iranian or Indian." Moreover, he suggested in his etymological vocabulary (II·143) that the base *i* was perhaps originally the Dravidian demonstrative base **i/ī*.

However, there is geographical and historical evidence to back up and make plausible our linguistically based historical judgment about Brahui. The *LSI* IV. 619 says of the Brahuis and their position in Baluchistan: " The Persians must, however, have invaded the country at a very early date. The cultivating class in the middle and westerly parts of the Khanat of Kalat [i.e. west of Brahui territory] are at the present day the Tajiks whose mother-tongue is Persian...... The last settlers were the Balochi who came from the south-west. They were not able to dislodge the Brahuis from the mountains, and they therefore took possession of the north-east and of the tract between Sind and Kacch Gandava." Balochi is in fact unanimously declared by Iranists to be not an eastern Iranian language, but to belong to the more westerly group of languages of the family. The *Grundriss der iranischen Philologie* 1.2.233, records that about 1600 A. D. the Hindu kings who still ruled in Kalat invited the Baloches from a more westerly area to come to their aid against the Afghans; Bray II.3 makes the Baloches come slowly from the west between the 7th and the 15th centuries. We may conclude that the Brahuis have only in the last centuries been in close contact with Baloches, that they may conceivably have had Persian-speaking neighbors earlier, but that for the earlier period and presumably for a long time they had more intimate contact with Indo-Aryan speakers (note the Hindu rulers of Kalat). These in all geographical probability were Sindhi speakers, and we should add that even if Sindhi speakers should have been absent from the Brahuis' present home, there is an evidently long-standing practice of the Brahuis whereby during the winter they migrate in large numbers into Sind where they have hereditary winter-quarters (cf. Bray II. 10–2). This in itself provides the bilingual situation which would allow an Indo-Aryan,

specifically a Sindhi, trait to diffuse into the Dravidian Brahui language.[29]

It is unfortunate that there are no general criteria that will take us any further in speculation on the history of the Brahui language. How long a period is required for a diffusion trait of this nature to gain such a foothold as it evidently has in Brahui ? We do not know. One point about the Brahui phenomena that bears on the history is that the 3rd personal forms are firmly established in all dialects of which Bray tells us anything, the 1st and 2nd sg. are most frequent in the southern area (Jhalawan), and the 1st and 2nd pl. are found in some dialects, the only one specified being Nichari, a Jhalawan dialect. This perhaps gives a clue to the direction of entry of the trait into the Brahui territory, i.e. we may guess that it gained entry first in the south, that the 3rd personal forms that occur all through the language were the first to develop- that these forms spread from the south all over the Brahui area, and that the other forms. developing later, have not spread to the same extent. But this is highly speculative

---

[29][Further study of Brahui has demonstrated that it has some structural traits which are clearly borrowed from Balochi and cannot have been borrowed from Indo-Aryan. It might be suggested that the trait treated in the text was also borrowed from Balochi, and that Brahui independently simplified the structure in a manner that parallels the Indo-Aryan structure. The only argument that could back up the suggestion that there was borrowing from Indo-Aryan would be at least one other such borrowing clearly demonstrated. There may be one such or even more. Even so, of course, there could be no certainty about the present trait. For presentation of the whole question at greater length, see my *Brahui and Dravidian Comparative Grammar*, Chapter IV, and ' Bilingualism and structural borrowing ', *Proceedings of the American Philosophical Society* 106.430–42 (1962).]

and of course adds nothing to the absolute chronology. All that seems really clear is our first point, that there is some probability that Brahui was in close contact with Indo-Aryan, in all probability Sindhi, before it established intimate contact with Iranian.

When we revert to examination of the historical relationships of this trait in the Iranian and Indo-Aryan languages, there are several possible solutions. On the basis of distribution and relative complexity, as was said above, we would prima facie judge that the pan-Iranian, more complicated trait has been borrowed in a less Iomplicated form (lack of feature 4) into the border cndo-Aryan languages, which are the only modern Indo-Aryan languages that have this structural feature and according to our basic assumption are to be suspected of having undergone diffusion in this matter.

Unfortunately, as soon as the historical data are examined, other possibilities present themselves. Both she branches of Indo-Iranian are well-known to have thown in their old and middle stages enclitic pronouns, which are derived at least in part from Indo-European (cf. the enclitic pronominal forms of Greek and of Hittite). The series for genitive and dative is as follows: Old Iranian (Old Persian, Avestan [Gatha/Younger]) 1sg. *-maiy*, *-mōi*/*-mē*; 2sg. *-taiy*, *-tōi*/*-ṯē*; 3sg. pl. *-śaiy*, *-hōi*/*-hē*, *-śē*; 1pl. ?, *-nə̱*/*-nō*; 2pl. ?, *-vō*/*və̱*. The corresponding Sanskrit series is: ***mē***, *tē*, —, *nas*/*ḥ*, *vas*/*ḥ*, to which Prakrit and Buddhist Hybrid Sanskrit add 3sg. pl. *sē* (cf. Hittite *-śe* < **-soy*) These series are in fact the origin of many of the pronominal suffixes. Other forms have entered into the origin of many of the modern forms, e.g. in the 1pl. *asma-* forms and in the 1sg. the verb form *asmi* 'I am.' And for some of the forms no origin has yet been found, e.g. the 3pl. *-kh* of Kashmiri. Since enclitic pronouns were commonly used in both branches of the Indo-Iranian family, it is possible that

in both branches the modern languages have their pronominal suffixes by inheritance and not by borrowing at all. This possibility can only be proved or disproved on phonological grounds.

The critical questions are: could all of the Indo-Aryan forms be borrowed from Iranian? Or are there any of the Indo-Aryan forms that could not be derived from Iranian forms, but can be derived from old Indo-Aryan forms? Some of the forms are in fact non-diagnostic, since on the two sides of the boundary some original forms have led by perfectly regular development to identical modern forms. Among these are *m* in the 1sg., which in both modern Iranian and the modern Indo. Aryan languages in question (though not in all modern Indo-Aryan languages) are the continuants of old *m*-forms- Similarly for *n*-forms in the 1pl., and *v*/*w*-forms in the 2pl. In 2sg. modern Iranian has some forms which preserve *t* and others which are supposed to have regularly lost *t* (so Morgenstierne for Parachi); in Indo-Aryan presumably *t* between vowels has been lost regularly at the Middle Indo-Aryan stage in the Sindhi and Lahnda developments.

The 3sg. *s*-forms of Indo-Aryan are almost certainly to be regarded as not borrowed from Iranian. Retention of *s* of *sē* is to be expected in Indo-Aryan. In Iranian the only language examined that has a sibilant is Persian, in which -*ś* is clearly derived from the *ś*-form of Old Persian and one of the Younger Avestan forms. Old Iranian generalized two treatments of an original *s*-form, one with the *ś* that appeared after *i* or *u* at the end of the preceding word, the other with the *h* that appeared after other vowels and after most consonants. All the other Iranian languages have in the 3sg. forms that can be derived from the *h*-form of the pronoun (though it is only Persian, with its retention of the other development, that gives the clue). Since it was not Persian that was contiguous

with the Indo-Aryan border languages, the *s*–forms of Indo-Aryan cannot be interpreted as anything but continuants of Middle Indo-Aryan forms.

The Indo-Aryan *s*–forms in the 1sg. (Sindhi, Lahnda, Kashmiri) and in the 1pl. (Sindhi, Lahnda) are to be derived from Indo-Aryan *–sm–* (1sg. *asmi*, 1pl. *asma–*), which yields modern *s* in the northwest, and otherwise *m*(*h*) (this statement of the development is over-simple, but will do in this context). This sibilant development cannot be related to the Iranian development of *m* (i.e. *–sm–* > *–hm–* > *–m–*), which is certainly seen in some of the Iranian 1pl. forms.

We find then at least three sets of forms which by their undeniably Indo-Aryan types of development are evidence that the modern Indo-Aryan forms are continuants of old Indo-Aryan forms, and which make it impossible to consider that the Indo-Aryan border languages have borrowed these particular forms from Iranian. From these particular forms we can legitimately make the extension to the whole set of suffixes and deny that Indo-Aryan has borrowed them from Iranian.

Such detailed examination as one can make of the Indo-Aryan languages in question, considering how defective the evidence is for most of the Dardic languages, arouses a suspicion that, though the Sindhi and Lahnda phonological shapes of the forms can be closely connected, it is impossible to reconstruct the details of the system for the whole of the northwestern group of languages. This suspicion is backed by the seeming fact that the structural feature (5) is only Dardic, and not Sindhi and Lahnda.

What then remains to be said? This much. Even though it seems indisputable that the Indo-Aryan frontier languages have inherited the pronominal suffixes from old

and middle Indo-Aryan, the rest of the modern Indo-Aryan languages have no structure of the kind that we have found, and there is a suspicion that there was no unified northwestern Indo-Aryan structure either. Can we not then ascribe the developments in the Indo-Aryan frontier languages to some extent to their Iranian neighbors? In other words, I would find here diffusion of the kind outlined in our primary assumption. The diffusion, however, is not one of details, but of a structure. Moreover, there was diffusion into languages which had already inherited a looser structure of the same kind, inchoate and somewhat meagrely structured, perhaps on the way to disuse as in the other Indo-Aryan languages. It was the influence, through bilingualism, of the strongly structured Iranian system that led to the complex and also strong, but somewhat disparate, structures that are seen today in the northwestern Indo-Aryan languages. And in addition, as we have found, the contiguous Dravidian language, Brahui, was also drawn, through bilingualism with the contiguous Indo-Aryan language, Sindhi, into the linguistic area contained by this isogloss that represents the pronominal-suffix structural trait. The diffusion was that of a structural feature, which was clothed with native morphemes in several different languages or language groups, and the direction of diffusion was Iranian to Indo-Aryan to Dravidian.

Should we extend the isogloss still further to include the Semitic languages? These show the same trait, with features (2) and (3), but not the specifically Iranian feature (4). Moreover, over several millennia there has been contact (and bilingualism) between Iranian and the Semitic languages, Mesopotamian, Aramaic, Arabic. There is good reason that we should speculate whether the Semitic trait may not have been operative in preserving in Iranian a trait that certainly flourishes there but that has died out in the sister Indo-Aryan except for the border languages.

We may neglect the Turkic languages, which show of the trait only feature (3), i.e. possessive suffixes with the noun (combined to be sure with feature 5). The comparatively modern chronology of contact between these languages and the southern Iranian–Semitic bloc makes it implausible on the whole to find Turkic either an influence or influenced.

We may also ignore Modern Greek, and the Romance languages, Italian, Iberian, and even French. These languages have a pronominal structure which shows some features in common with those we have been discussing. Historically the Modern Greek system derives from the classical Greek enclitics. The forms of the Romance languages are derived from Latin forms which are not enclitic; i.e. the Romance systems are historically not derived from the older Indo-European enclitic system The geographical area is so discontinuous with the center of our interest that we may feel safe in concluding that there can be no direct historical connection between the Mediterranean languages and those of the Indo-Iranian group under consideration, except in so far as the Greek system and the old Indo-Iranian system both derive from the older Indo-European system.

Nor need we do more than mention the enclitic pronouns of Hittite. Their striking variant of feature (4), viz. the bunching of enclitics, pronominal and others, immediately after the first word of the sentence, throws much light on early Greek and Vedic Sanskrit features. But what Hittite has in common with anything that has been found in modern Iranian and northwestern Indo-Aryan, it has in common only because it probably represents fairly closely the earliest stage of Indo-European of which we have any knowledge, and because these modern forms of Indo-Iranian turn out to have preserved for more than four millennia more of the old Indo-European system of pronominal enclitics than any other of the contemporary Indo-European languages except Modern Greek.

Another structural trait which lends itself to the same type of treatment may be handled more concisely. This feature is grammatical gender. It is found in Indo-Aryan is general, and, although it is lost in modern Magadhan, it remains in some form almost everywhere else. The northwest retains only two genders (masculine and feminine) like most of the rest of Indo-Aryan with the exception of Marathi and Gujarati, which retain all three genders of Old and Middle Indo-Aryan. In Kashmiri (as well as Shina and some of the Kohistani languages) and Sindhi, nouns, adjectives, demonstrative pronouns, and some parts of the verb system show the gender structure. In Lahnda, however, the demonstrative pronouns have no gender-differentiated forms, even though a gender structure exists elsewhere in the language. Torwali, and of the Kafir group, Waigali, Pashai, Bashgali and Gawar-bati, have systems similar to that of Lahnda; Khowar is said by the *LSI* not to have gender at all.

West Iranian in general has retained little of the Old Iranian gender system, which was the same as that of Old Indo-Aryan, although there are traces, which Morgenstierne summarizes in *Handbuch der Orientalistik*, p. 160. Such languages as standard Kurdish agree with Persian in having none, as does also Balochi. In East Iranian Wakhi and Sanglechi have completely lost gender; so has one dialect of Ormuri, although the other preserves a two-gender system in the adjectives, demonstratives, and parts of the verb structure. Traces of such a system are found in some of the northeastern Pamir languages (see Morgenstierne, *loc. cit.*). Pashto also has this sort of two-gender system; the distinction, however, is not overt in the direct (non-oblique) forms of the demonstrative pronouns.

The Dravidian language Brahui belongs to a family which has a system of natural gender in the demonstrative pronouns and by agreement in some parts of the verb

structure in some of the languages. Two of the languages, Toda and Brahui, have lost all traces of this system, and both of them in the demonstrative pronouns have retained as the now general form the old non-personal form, e.g. Toda aθ 'he, she, it' (plural aθa·m) < *atu. The Brahui dād 'this person or thing' is related to *atu, as is the plural dāfk to *avai (i.e. -ād : *at-, -āf-: *av-); the problem of the phonemic shape of the stem and the meaning in Brahui has been discussed elsewhere. What is obvious and relevant here, is that in form the Toda and Brahui pronouns show difference in detail in their development. Their history must be independent, and we may conclude that the loss of the gender distinctions is independent. There seems to be nothing at all in the languages immediately contiguous to Toda that is parallel to its development; the conclusion must be that the loss of gender took place in Toda without any stimulus from outside. Shall we say the same of Brahui? Undoubtedly not, for it is in an area where loss of the old gender has been widespread.

Both sides of the Indo-Iranian border are complicated in the distribution of gender systems, and no simple historical solutions can be given. On the Iranian side, the spotty distribution of two-gender systems, which we know to be archaic as compared with the lack of a gender system in most of the languages, would lead to an interpretation of the gender-distinguished languages as showing retention of an archaism in isolated pockets or at the margins, and loss elsewhere under the influence of certain languages of widespread prestige, such as Persian (so, e. g.. Morgenstierne has said about that one of the Ormuri dialects which has no gender). On the other hand, the similarities that appear between the Indo-Aryan languages and the Iranian languages that are contiguous to one another on the two sides of the border make it imperative to view the matter in the light of the general principles which we have stated on diffusion.

Specifically, it seems suspicious that there is a bunching of Iranian languages with two-gender systems in the area nearest to the Indo-Aryan border, on the other side of which all the languages (with the exception of Khowar) have two-gender Indo-Aryan systems. It is tempting to say that the rate of retention of the archaic Iranian system is higher where there is influence from Indo-Aryan. Balochi, notably, which is said to be a West Iranian language of comparatively recent arrival in its present geographical situation, is of Persian type in this matter, and presumably arrived in the border area after it had lost gender and therefore too late to undergo the Indo-Aryan influence. Khowar, on the other side, if it is correctly reported to have no gender, may well owe this trait to Iranian Wakhi and Sanglechi to the north; these languages in all probability owe their loss to Persian influence. In this matter of the incidence of gender-systems, then, those East Iranian languages that have them have certainly preserved an Iranian archaism, but there is strong suspicion that the retention owes something to Indo-Aryan influence; moreover, at least one Indo-Aryan language, Khowar, seems to have been influenced by Iranian. I have assumed here that phonological and morphological details, when worked out, will not contradict these statements, which are based on distribution alone; this is not the place. nor have I the competence in Iranian, to examine the details.

One set of details, however, very obviously yields further statements, which have a high degree of probability. These have to do with the demonstrative pronouns. On the Indo-Aryan side, Lahnda, Torwali, and some of the Kafir languages agree in having a gender system and also in having no gender-differentiated forms in the demonstrative pronouns. The Iranian Pashto also has a gender system which shows up even in the demonstrative pronouns, but the direct (non-oblique, essentially subject) forms have no gender distinction. The map

shows these Indo-Aryan languages strung out along the northeastern and eastern border of the Pashto area, and moreover at the southern end of the Lahnda area a wedge of Balochi territory (with no gender at all) between Lahnda and Pashto. These Indo-Aryan languages are only in part contiguous with one another; Torwali abuts on Lahnda, but Khowar cuts off the Kafir languages from Torwali. There is a possibility, then, that it is the highly specific, skew feature in the demonstrative pronouns of Pashto, which certainly cannot be separated from the Iranian tendency towards loss of gender, that influenced the contiguous Indo-Aryan languages, perhaps in part independently, towards a general, non-skewed loss of gender in the demonstrative pronouns, which shows up nowhere else in that great area of Indo-Aryan territory in which gender is preserved. This loss of gender is of Iranian type; prima facie one would argue Iranian influence, and yet there cannot be complete certainty.

Since Brahui shows no trace at all of a gender system, Iranian influence can be postulated, Balochi or Persian (but certainly not Pashto). There is no need for recourse to primary or secondary Indo-Aryan influence such as we have found necessary elsewhere in historical interpretation of Brahui features. In any case, Lahnda could have had very little influence on Brahui developments, and Sindhi which has had considerable influence, cannot be invoked here since it retains a full-fledged two-gender system of the normal Indo-Aryan type.[30]

---

[30]Iranian influence as causing the loss of the gender system in Brahui has already been pointed out in several places, e. g. *LSI* IV. 622, and Jules Bloch, *Structure grammaticale des langues dravidiennes* (1946), 5, but without any detail, and without knowledge of the Toda phenomena.

Bibliography

Linguistic Survey of India, Vol. IV, Muṇḍā and Dravidian languages (1906); Vol. VIII. 1, Indo-Aryan family, north-western group, specimens of Sindhī and Lahndā (1919); vol. VIII. 2, specimens of the Dardic or Piśācha languages (including Kāshmiri) (1919); vol. X, specimens of languages of the Eranian family (1921).

Wilhelm Geiger and Ernst Kuhn, *Grundriss der iranischen Philologie*, 1er Bd., Strassburg, 1895–1901.

Handbuch der Orientalistik, 1te Abt. Der nahe und der mittlere Osten, 4ter Bd. Iranistik, 1ter Absch. Linguistik, Leiden, 1958.

T. Grahame Bailey, *Grammar of the Shina* (*Ṣiṇā*) *Language* (Royal Asiatic Society, Prize Publication Fund, VIII), London, 1924.

George Waters Gilbertson, *The Balochi Language*, Hertford, 1923.

George A. Grierson, *The Piśāca Languages of North-western India.* (Asiatic Society Monographs, VIII), London, 1906.

id., *Ishkashmi, Zebaki, and Yazghulami* (Royal Asiatic Society, Prize Publication Fund, V), London, 1920.

id., *Torwali, an Account of a Dardic Language of the Swat Kohistan* (Royal Asiatic Society, Prize Publication Fund, XI), London, 1929.

Ann K. S. Lambton, *Three Persian Dialects* (Royal Asiatic Society, James G. Forlong Fund, XVII), London, 1938.

Gilbert Lazard, *Grammaire du persan contemporain*, Paris, 1957.

D. L. R. Lorimer, *A Syntax of Colloquial Pashtu*, Oxford, 1915.

id., *The Wakhi Language*, London, 1958.

Ernest N. McCarus, *A Kurdish Grammar* (American Council of Learned Societies, Program in Oriental Languages, Ser. B, 10), New York, 1958.

Georg Morgenstierne, *Report on a Linguistic Mission to Afghanistan* (Instituttet for sammenlignende Kulturforskning, Ser. C I–2) Oslo, 1926.

id., *Indo-Iranian Frontier Languages*, vol. I, Parachi and Ormuri (*ibid.*, Ser. B XI), Oslo, 1929.

id., *Report on a Linguistic Mission to North-western India* (*ibid.*, Ser. C III–1), Oslo, 1932.

id., *Indo-Iranian Frontier Languages*, vol. II, Iranian Pamir Languages (*ibid.*, Ser. B XXXV), Oslo, 1938.

id., *Indo-Iranian Frontier Languages*, vol. III, The Pashai Language, 2. Texts and translations (*ibid.*, Ser. B XL), Oslo, 1944.

id., 'The language of the Ashkun Kafirs,' *Norsk Tidsskrift for Sprogvidenskap* (*NTS*) 2·192–289 (1929).

id., 'Notes on Shumashti,' *ibid.*, 13·239–81 (1945).

id., 'The Waigali language,' *ibid.*, 17·146–324 (1954)

Herbert Penzl, *A Grammar of Pashto* (American Council of Learned Societies, Program in Oriental Languages, Ser. B. 2), Washington, D. C., 1955.

Denys DeS. Bray, *The Brahui Language*, I. Introduction and grammar, Calcutta, 1909.

id., *The Brāhūī Language*, II. The Brāhūī Problem, III. Etymological vocabulary, Delhi, 1934.

T. Burrow and M. B. Emeneau, *A Dravidian Etymological Dictionary*, Oxford, 1961.

Feature	Persian	Kurdish	Parachi and Ormuri	Shighni	Ishkashmi	Wakhi	Pashtu	Balochi	Brahui	Sindhi	Lahnda	Kashmiri	Shina	Shumashti	Pashai	Bashgali (Kati)	Waigali	Ashkun
(1)	old mod. x > o	x	x	x	x	x	o	?	o	x	x	x	x	x	x	x	?	?
(2)	x	x	x	?	?	o	x	x	x	x	x	x	x	x	x	x	x	?
(3)	x	x	x	?	?	o	x	x	x	x	x	o	o	x	x	x	x	x
(4)	x	x	x	x	x	x	x	x	o	o	o	o	?	o	o	o	o	o
(5)	x	x	x	?	x	x	o	o	o	o	o	x	?	x	x	x	x	x
(6)	o x mod. dial	x	o	?	?	o	x	o	o	o	o	o	?	o	o	o	o	o
									Special features									

Table 1. Distribution by languages of the features characterizing the pronominal suffixes.

	Persian	Kurdish	Parachi	Ormuri	Shighni	Zebaki	Wakhi	Pashtu	Balochi	Brahui	Sindhi	Lahnda	Kashmiri	Shumashti	Bashgali	Waigali	Ashkun
1sg.	–(a)m	–(ị)m	–(u)m	–(a)m,	–am, –um	–am –em, –im	–ɛm, –əm	mi	–u^{n}	–ka	–s^{e}, –m^{e}	–s, –m	–s, –m	–m	–m, –ī	–m	–a
2sg.	–(e)t	–(i)t	–a(u)	–(a)t	–at	–ē, –ī, –ai	–ɛt, –ət	di	–it	–ne	–ĕ^{n}, –^{e}	–ē^{n}, etc.	–kh –th –y	–u	–sh	–sh	–ag
3sg.	–(e)š	–i, –y	–ē	–a, –wa	–i, –e	–a	–i	(y)ee	–ē, –ī	–ta(n)	—, –s^{e}, –āī^{n}	—, –s	—, –n, –s	–s	–s, –as	–s, –sha	–(a)s
1pl.	–(e)mun, (e)mân	–man	–(a)n	–ēn	–am	–en –e	–ɛn, –ən	mu	–ü^{n}	–nan	–sī^{n}, –sū^{n}, –ū^{n}	–s, –sē, –ahsē	—	–ani/e	–mish	–mish	–kəṇä, –kə^{n}ṛä
2pl.	–(e)tun, –(e)tân	–tan	–ō(u)	?	–et	–ev, –av	–ɛv, –əv	mu	–it	–num	–^{u}, –w^{a}	—, –nē, etc.	–wa	–i	–^{n}r	–e^{n}	–aṇeg, –a^{n}ṛeg
3pl.	–(e)šun, –(e)šân	–yan	–an, –ē	–a, –wa	–en	–en, –e	–ɛv, –əv	(y)ee	–ish	–tā(n)	—. ·n^{e},–n^{a}, –āū^{n}	—, –nē	—, –kh	–at	?	–t	–(a)ṇes, –(a^{n})ṛes

Table 2 Pronominal suffixes.